MECÁNICA CLÁSICA

Rafael Pérez Pascual

29 de enero de 2008

Instituto de física, UNAM.

Índice general

1. Introducción **4**
- 1.1. Ecuaciones diferenciales en el espacio fase. 4
- 1.2. Ecuaciones diferenciales lineales. 8
- 1.3. Ecuaciones diferenciales no lineales. 19
- 1.4. El concepto de flujo en el espacio fase. 22

2. La mecánica vectorial. **25**
- 2.1. Sistemas mecánicos en una dimensión. 27
  - 2.1.1. El oscilador armónico. 32
  - 2.1.2. El péndulo simple. 33
- 2.2. El movimiento en dos o más dimensiones. 38
  - 2.2.1. Figuras de Lissajous. 40
  - 2.2.2. El campo central. 41
  - 2.2.3. El péndulo esférico. 53
  - 2.2.4. El péndulo doble. 59
  - 2.2.5. El cuerpo rígido libre. 61

3. Formulación lagrangiana de la mecánica. **73**
- 3.1. Espacio de configuración, coordenadas generalizadas y ecuaciones de movimiento. 74
- 3.2. Constricciones. 92
- 3.3. Mecánica lagrangiana de sistemas holonómicos. 97
- 3.4. Coordenadas ignorables. 106
- 3.5. Integrales de movimiento para lagrangianas independientes del tiempo y el teorema del virial. 115
- 3.6. Las fuerzas de constricción. 118
- 3.7. Sistemas no holonómicos. 124
- 3.8. El principio de Hamilton. 135
- 3.9. Simetrías e integrales de movimiento. 141

3.10. Variedades diferenciales. 148

4. Oscilaciones pequeñas. 167
4.1. Modos normales de oscilación. 167
4.2. Medios continuos. 177

5. Formulación hamiltoniana de la mecánica. 185
5.1. Las ecuaciones de Hamilton. 186
5.2. Los teoremas de Liuville y Poincaré. 192
5.3. EL espacio fase. 197
5.4. El flujo hamiltoniano. 221
5.5. El teorema de Noether en la formulación hamiltoniana. 228
5.6. Transformaciones canónicas. 243
5.6.1. Transformaciones canónicas independientes del tiempo. 247
5.6.2. Transformaciones canónicas dependientes del tiempo. . 257

6. Formulación de Hamilton Jacobi. 264
6.1. La ecuación de Hamilton Jacobi. 264
6.2. Ejemplos del proceso de Hamilton Jacobi. 269
6.3. La aproximación de la eikonal y la ecuación de Hamilton Jacobi.285
6.4. La función principal de Hamilton. 294
6.5. La ecuación de Hamilton Jacobi y transformaciones canónicas. 303

7. Sistemas mecánicos integrables. 305
7.1. Sistemas integrables. 305
7.2. Variables de acción y ángulo. 320
7.3. Invariantes adiabáticos. 332

8. Sistemas mecánicos no integrables. 338
8.1. Existencia de los sistemas no integrables. 339
8.2. Las trayectorias en los sistemas no integrables. 344
8.3. EL teorema de KAM. 356
8.4. Caos hamiltoniano. 382
8.5. Ejemplos ilustrativos de sistemas no integrables. 406
8.5.1. El rotor pateado. 406
8.5.2. La hamiltoniana de Henon y Heiles. 422
8.5.3. Billares. 426
8.5.4. El péndulo doble. 435
8.6. Caracterización del caos. 448

Capítulo 1

Introducción

En la base de la mecánica clásica están las ecuaciones diferenciales. Podemos decir que Newton inventó el cálculo y las ecuaciones diferenciales para poder expresar sus ideas sobre el movimiento de los cuerpos en forma matemática. La segunda ley de Newton es una expresión, en forma de ecuación diferencial, de la relación física entre la aceleración, la masa y la fuerza. Por tanto los conceptos entorno a las ecuaciones diferenciales tienen un papel importantísimo en la mecánica clásica y en general en toda la física. Haremos una breve revisión de la teoría de las ecuaciones diferenciales con el objeto de fijar conceptos, resultados y notaciones que serán útiles a lo largo de todo el estudio de la mecánica.

1.1. Ecuaciones diferenciales en el espacio fase.

Una ecuación diferencial ordinaria de orden n se expresa en forma tradicional como

$$f\left(\frac{d^n x}{dt^n}, \frac{d^{n-1}x}{dt^{n-1}}, \cdots, \frac{dx}{dt}, x, t\right) = 0, \tag{1.1}$$

donde $f : R^{n+1} \times R \to R$ es una función de clase C^1.[1] Una solución será una función $x : R \to R$, expresada como $x(t)$, tal que al ser sustituida en la ecuación (1.1), ésta se satisfaga. El teorema de existencia y unicidad nos dice que, dadas unas condiciones iniciales, esto es un conjunto de valores específicos de la función $x(t)$ y de sus derivadas a un cierto tiempo t_0, existe

[1] Para el objeto de este texto simplemente supondremos que es suficientemente derivable como para que sean validos los teoremas de existencia y unicidad y que las ecuaciones tengan sentido físico.

una y sólo una solución de la ecuación diferencial que, además de satisfacer la ecuación, cumple con dichas condiciones iniciales.

Veamos las ecuaciones diferenciales desde otro punto de vista.

Tomemos nuevas variables definidas por

$$\begin{aligned} y_1 &= x, \\ y_2 &= \frac{dx}{dt}, \\ &\vdots \\ y_n &= \frac{d^{n-1}x}{dt^{n-1}}. \end{aligned} \tag{1.2}$$

Notamos que podemos poner

$$\begin{aligned} \frac{dy_1}{dt} &= y_2, \\ \frac{dy_2}{dt} &= y_3, \end{aligned} \tag{1.3}$$

$$\begin{aligned} &\vdots \\ \frac{dy_{n-1}}{dt} &= y_{n,} \end{aligned} \tag{1.4}$$

$$\frac{dy_n}{dt} = \frac{d^n x}{dt^n},$$

por tanto podemos expresar la ecuación diferencial en la forma

$$f\left(\frac{dy_n}{dt}, y_n, \cdots, y_2, y_1, t\right) = 0 \tag{1.5}$$

o despejando

$$\frac{dy_n}{dt} = g(y_1, \cdots, y_n, t). \tag{1.6}$$

Vemos que, denotando la derivada con respecto al tiempo por medio de un punto encima de la variable, la ecuación diferencial de orden n es totalmente equivalente al sistema de n ecuaciones diferenciales de primer orden acopladas

$$\begin{aligned} \dot{y}_1 &= y_2, \\ \dot{y}_2 &= y_3, \\ &\vdots \\ \dot{y}_n &= g(y_1, \cdots, y_n, t). \end{aligned} \tag{1.7}$$

Inversamente es fácil ver que si tenemos un sistema de n ecuaciones diferenciales de primer orden acopladas

$$\begin{aligned} \dot{y}_1 &= h_1(y_1, y_2, \cdots, y_n, t), \\ \dot{y}_2 &= h_2(y_1, y_2, \cdots, y_n, t), \\ &\vdots \\ \dot{y}_n &= h_n(y_1, y_2, \cdots, y_n, t), \end{aligned} \tag{1.8}$$

con las funciones h formando un conjunto linearmente independiente, será equivalente a una sóla ecuación diferencial de orden n. Podemos por tanto contemplar a las ecuaciones diferenciales ordinarias como una ecuación de la forma

$$\dot{y} = h(y, t), \tag{1.9}$$

donde

$$y = (y_1, y_2, \cdots, y_n) \tag{1.10}$$

es un punto de R^n y

$$h : R^n \times R \to R^n \tag{1.11}$$

una función de y y t en R^n.

Cuando una ecuación diferencial no depende explícitamente del tiempo, esto es cuando la función h no contiene al argumento t, decimos que se trata de una ecuación autónoma, en caso contrario de una no autónoma. En esta exposición resumida de la teoría de las ecuaciones diferenciales nos restringiremos a las ecuaciones autónomas, no sin antes hacer notar que la perdida de generalidad no es grande, ya que una ecuación no autónoma de orden n es equivalente a una autónoma de orden $n + 1$. Así de ahora en adelante trataremos únicamente con ecuaciones autónomas.

Resumiendo, una ecuación diferencial consta de un espacio R^n al que se le llama espacio fase[2] y una expresión de la forma

$$\dot{y} = h(y), \tag{1.12}$$

donde $y \in R^n$ es un punto del espacio fase y $h : R^n \to R^n$ es un mapeo del espacio fase en sí mismo. Por una solución entenderemos un mapeo de los reales, que representan al tiempo, en el espacio fase

$$y : R \to R^n, \tag{1.13}$$

[2]Este concepto de espacio fase como el espacio en el que se define una ecuación diferencial, difiere del de espacio fase en el ámbito de la mecánica hamiltoniana. Cuando veamos este último concepto se verá con claridad la diferencia, también veremos que el contexto deja claro a cual nos referimos y que por tanto, no hay temor a confusiones.

que satisfaga la ecuación y que cumpla con las condiciones iniciales

$$y(t_0) = y_0. \tag{1.14}$$

Desde una perspectiva geométrica la solución (1.13) es una curva en el espacio fase parametrizada por el tiempo.

El teorema de existencia y unicidad nos hace ver que para cada punto $y_0 \in R^n$ existe una y sólo una solución de la ecuación diferencial que tiene a ese punto como condición inicial, esto es, por cada punto del espacio fase pasará una curva solución, por lo que estas curvas llenan el espacio fase y, como la soluciones son únicas, no se intersectan, esto es, por cada punto pasa una y sólo una solución. Esta forma de ver las ecuaciones diferenciales tiene muchas ventajas sobre la representada por medio de la ecuación (1.1), veamos algunas de ellas.

En primer lugar se construye un espacio más amplio donde trabajar, el espacio fase. En él requerimos utilizar únicamente primeras derivadas en vez de las de órdenes superiores. La solución de la ecuación diferencial se ve como un mapeo de R en R^n o sea como una curva en el espacio fase parametrizada por el tiempo. El concepto de solución adquiere un sentido geométrico: la solución, para una condición inicial dada, es una curva en el espacio fase sobre la que se satisface la ecuación y que pasa por el punto que representa a la condición inicial.

En segundo lugar el mapeo h asocia a cada punto de R^n otro punto o vector de R^n, o sea podemos entender a h como un campo vectorial en el espacio fase, por decirlo pictóricamente como una flechita anclada en cada punto del espacio fase. La ecuación diferencial indica que la derivada con respecto al tiempo de una curva solución en cualquiera de sus puntos, o sea la velocidad con la que se recorre la curva o mejor dicho el vector tangente a la curva en el punto, debe coincidir con el vector del campo vectorial en ese mismo punto. Una solución, por tanto, no es otra cosa que una curva tangente, en todos sus puntos, al campo vectorial definido por la ecuación diferencial. Desde este punto de vista geométrico, el teorema de existencia y unicidad nos dice que dado un campo vectorial de clase C^1 en un espacio R^n, existe un conjunto de curvas tangentes al campo vectorial que llenan dicho espacio, esto es, que por cualquier punto de dicho espacio pasa una de ellas y que lo hacen de tal forma que no se intersectan, pues de hacerlo por el punto de intersección pasarían dos de esas curvas o sea habría en ese punto dos vectores del campo vectorial.

Pensemos en un sistema físico caracterizado totalmente por n variables, o sea tal que su estado físico queda totalmente definido por esas n variables.

La evolución temporal de ese sistema quedará definida al especificar cada una de las n variables como función del tiempo y, si el sistema admite ser estudiado por medio de una ecuación diferencial, ésta puede expresarse en la forma (1.8). Vemos que el espacio fase vendría a ser el espacio de estados del sistema físico, la evolución del sistema vendría a ser una curva en ese espacio y la ecuación diferencial, que a fin de cuentas expresa la física del sistema, puede ser interpretada como la tendencia lineal al cambio del sistema en cada uno de sus estados como función del estado mismo.

Tomemos como ejemplo al oscilador armónico. La ecuación de movimiento se obtiene de la segunda ley de Newton

$$m\frac{d^2x}{dt^2} = -kx. \tag{1.15}$$

Para pasar de esta forma a la dada por (1.8) definimos la velocidad por $v = dx/dt$ y entonces podemos escribir

$$\begin{aligned} \dot{x} &= v, \\ \dot{v} &= -\frac{k}{m}x. \end{aligned} \tag{1.16}$$

El espacio fase es el espacio R^2 compuesto por los puntos (x, v) y el campo vectorial asocia al punto (x, v) el vector $(v, -kx/m)$, esto es en cada punto del espacio fase o estado del sistema físico (x, v) hay una tendencia lineal al cambio en la que x tiende a cambiar con magnitud v y v con magnitud $-kx/m$. Gráficamente se vería como se esquematiza en la figura 1.1

Las soluciones son curvas cerradas concéntricas en torno al origen, sabemos que son elipses, como puede verse claramente al trazar las tangentes al campo vectorial. Estas curvas generan, al proyectarse en los ejes coordenados, la oscilación periódica en x y la correspondiente en v desfasada noventa grados.

1.2. Ecuaciones diferenciales lineales.

En la teoría de las ecuaciones diferenciales lineales con coeficientes constantes escribimos la ecuación en la forma

$$a_n\frac{d^nx}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0x = 0., \tag{1.17}$$

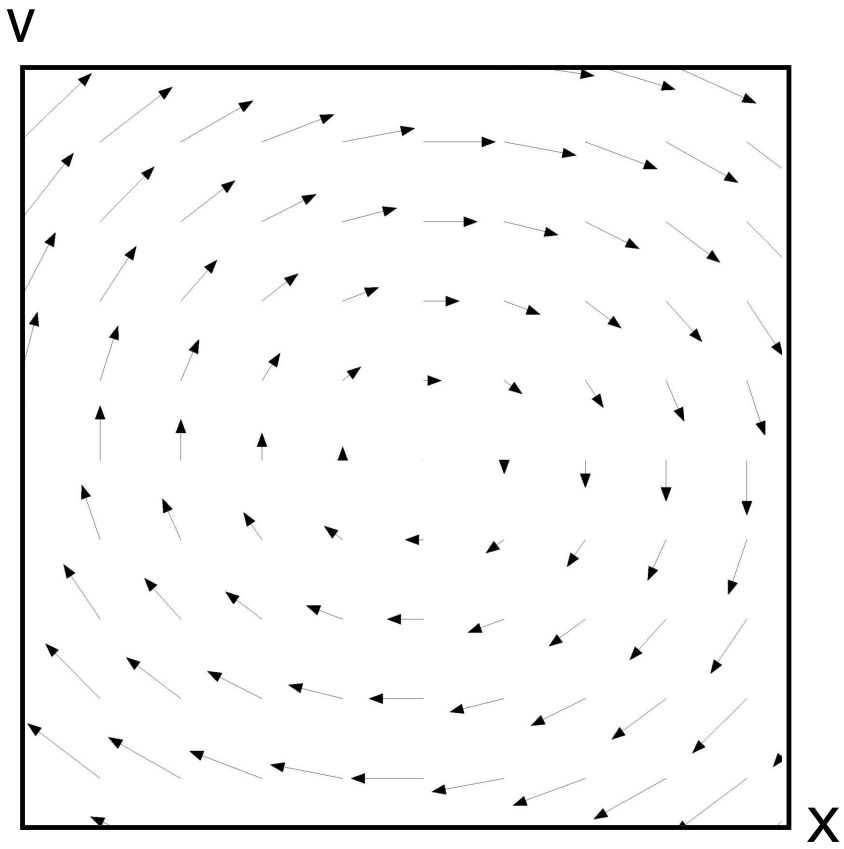


Figura 1.1: Campo vectorial del oscilador armónico en el espacio fase.

donde las a son constantes reales. Sabemos que las soluciones son combinaciones lineales de las exponenciales

$$e^{\lambda_i t}, \tag{1.18}$$

donde las lambdas son las raíces del polinomio que se forma al sustituir las derivadas en la ecuación (1.17) por una literal algebraica z

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0. \tag{1.19}$$

En el espacio fase la ecuación diferencial lineal se verá como

$$\frac{dy}{dt} = \mathcal{A} y, \tag{1.20}$$

donde y es un punto de R^n y $\mathcal{A}$ es una matriz real no singular de $n \times n$.

Hagamos un cambio lineal de coordenadas en el espacio fase para pasar de las coordenadas y a unas nuevas coordenadas u. Denotemos por $\mathcal{B}$ a la matriz no singular que representa dicho cambio de coordenadas, entonces

$$u = \mathcal{B} y \tag{1.21}$$

y

$$y = \mathcal{B}^{-1} u. \tag{1.22}$$

En las nuevas coordenadas la ecuación diferencial quedará expresada por

$$\frac{du}{dt} = \mathcal{B}\mathcal{A}\mathcal{B}^{-1} u. \tag{1.23}$$

Ahora si el cambio de coordenadas es tal que hace que la matriz $\mathcal{A}$ quede diagonalizada

$$\mathcal{B}\mathcal{A}\mathcal{B}^{-1} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda_n \end{pmatrix}, \tag{1.24}$$

donde las λ_i son los eigenvalores de $\mathcal{A}$, las ecuaciones que estaban acopladas se desacoplan y nos quedan n ecuaciones de primer orden en una sóla variable

$$\dot{u}_i = \lambda_i u_i \qquad i = 1, \cdots, n, \tag{1.25}$$

las que tienen por solución

$$u_i = C_i e^{\lambda_i t}, \tag{1.26}$$

donde las constantes C_i quedan determinadas por las condiciones iniciales. Vemos así que la ecuación (1.19) no es otra cosa que la ecuación característica de la matriz $\mathcal{A}$, y vemos que para resolver la ecuación diferencial no hicimos otra cosa que buscar un sistema de coordenadas en el cual la ecuación encontrara su expresión mas simple: aquella en que la matriz $\mathcal{A}$ queda diagonalizada y el sitema queda desacoplado.

En el caso en que la matriz $\mathcal{A}$ no sea diagonalizable, esto es que presente eigenvalores complejos, sabemos que es posible llevarla por medio de un cambio lineal de coordenadas a una forma diagonal en términos de bloques de dos por dos. Esto separa la ecuación original en bloques independientes de dos ecuaciones lineales acopladas en dos variables; pero sabemos que en este caso la solución del sistema de dos por dos es una combinación lineal de las exponenciales de sus dos eigenvalores complejos conjugados multiplicados por el tiempo y llegamos a algo similar al caso diagonalizable. Naturalmente los casos excepcionales en que la matriz tenga eigenvalores iguales hay que tratarlos aparte, pero no ofrecen dificultades conceptuales adicionales.

Vemos con este ejemplo una ventaja de formular las ecuaciones diferenciales en el espacio fase, nos ha permitido introducir directamente las herramientas y conceptos del álgebra lineal y adquirimos una visión más amplia del significado de la ecuación, de su solución y de los métodos empleados para resolverla.

Ésta manera de estudiar las ecuaciones diferenciales lineales nos abre la posibilidad de clasificar los posibles comportamientos generales de ellas de una manera muy simple. Comencemos por el caso de las ecuaciones en dos dimensiones.

Tomemos una ecuación de la forma $\dot{x} = \mathcal{A}x$ donde $x \in R^2$ y $\mathcal{A}$ es una matriz no singular de dos por dos. Notemos en primer lugar que el origen es, por si mismo, a la vez un punto del espacio fase y una solución de la ecuación diferencial, esto es la función de R en R^2 que asocia a todo tiempo el valor constante $(0, 0)$ es solución, ya que su derivada es cero, y al aplicar la matriz $\mathcal{A}$ al origen obtenemos también cero, por lo que se cumple la ecuación. En general, cuando un punto del espacio fase es solución de una ecuación diferencial por sí mismo, se le llama punto crítico y corresponde a soluciones estáticas o de equilibrio en las que no hay cambio del sistema con el paso del tiempo; así podemos decir que toda ecuación lineal tiene al origen como punto crítico o de equilibrio.

Las soluciones de la ecuación serán combinaciones lineales de las exponenciales del producto del tiempo por los eigenvalores de la matriz que la define, λ_1 y λ_2

$$x_1 = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}, \tag{1.27}$$

$$x_2 = c_3 e^{\lambda_1 t} + c_4 e^{\lambda_2 t}, \tag{1.28}$$

las constantes de la combinación lineal estarán determinadas por las condiciones iniciales y sólo dos de las cuatro que hemos señalado son independientes; en el caso del origen como solución o punto crítico estas constantes son cero y corresponden a tomar al origen como condición inicial.

Si los dos eigenvalores son reales habrá un sistema de coordenadas en el que la matriz se expresa en forma diagonal, denotando con y a este sistema de coordenadas, las soluciones adquirirán la forma

$$\begin{aligned} y_1 &= d_1 e^{\lambda_1 t}, \\ y_2 &= d_2 e^{\lambda_2 t}, \end{aligned} \tag{1.29}$$

donde d_1 y d_2 son constantes que representan, en ese sistema y de coordenadas, a las condiciones iniciales.

Vemos que las rectas correspondientes a los ejes coordenados y_1 y y_2 son de por sí soluciones de la ecuación, lo que ocurre cuando las condiciones iniciales hacen que d_1 o d_2 sean cero. Notamos que en realidad se trata de cinco soluciones, ya que los ejes coordenados se intersectan en el origen que es de por sí una solución y por tanto no puede formar parte de ninguna otra, así la parte positiva y la negativa de cada uno de los ejes coordenados quedan separadas como solución, por lo que tenemos cuatro, más el origen mismo, hacen las cinco que mencionamos.

Cuando el eigenvalor correspondiente sea positivo, la trayectoria, esto es la recta determinada por uno de los semiejes, se recorrerá con el transcurrir del tiempo en la dirección que se aleja del origen y tenderá asintóticamente al infinito; cuando el eigenvalor sea negativo la trayectoria se recorrerá en dirección al origen y tenderá asintóticamente a este. Es importante recordar que, en este último caso, si bien la trayectoria tiende al origen, nunca llega, pues de hacerlo tendríamos dos soluciones distintas en el mismo punto lo que violaría la unicidad de las soluciones; naturalmente en el caso de un eigenvalor positivo tendremos una situación similar para tiempos negativos.

En el caso general, esto es, cuando la condición inicial nos coloca en un punto del espacio fase en el que ni d_1 ni d_2 son cero, tendremos una combinación del comportamiento en cada uno de los ejes coordenados. Observamos que la trayectoria o solución debe permanecer en el mismo cuadrante, de lo contrario tendría que intersectar a uno de los ejes que es a su vez una trayectoria, en violación a la unicidad. Puede haber tres casos distintos: los

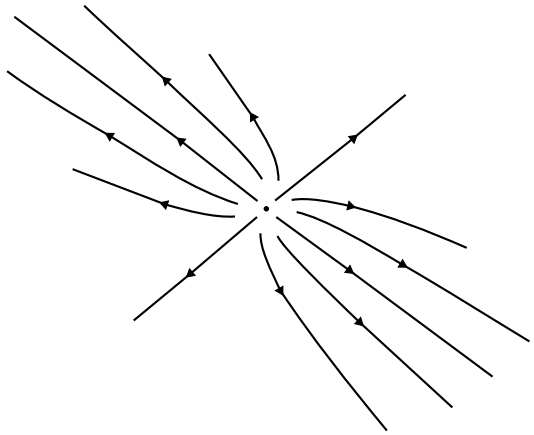


Figura 1.2: Un nodo inestable.

dos eigenvalores son positivos, los dos son negativos o uno es positivo y el otro negativo.

Cuando ambos eigenvalores son positivos las trayectorias se alejarán exponencialmente del origen tanto en la dirección de uno de los ejes coordenados como en la del otro, y si las vemos desde la perspectiva de los tiempos negativos las veremos tender asintóticamente al origen. Esto produce un diagrama de soluciones en el espacio fase como el que se muestra en la figura 1.2

Si los eigenvalores resultan ser ambos negativos tendremos un caso similar al anterior pero las trayectorias se recorren en el sentido inverso como puede verse en la figura 1.3

En el caso en que uno de los eigenvalores sea positivo y el otro sea negativo, una de las coordenadas tenderá a infinito, mientras que la otra tenderá a cero. Esto nos conduce a un comportamiento hiperbólico en el que los ejes coordenados son las asíntotas de unas curvas hiperbólicas que representan las trayectorias o soluciones de la ecuación diferencial, como se muestra en la figura 1.4

Estas figuras ilustran, de manera cualitativa, el comportamiento general de las trayectorias en el espacio fase de las ecuaciones diferenciales lineales en dos dimensiones cuando los dos eigenvalores de la matriz que define a la ecuación son reales; naturalmente la forma exacta de las curvas y la rapidez con que se recorren con el paso del tiempo dependerán de los valores de los eigenvalores de la ecuación, pero no habrá comportamientos distintos a estos desde el punto de vista cualitativo o topológico. Nótese que el comporta-

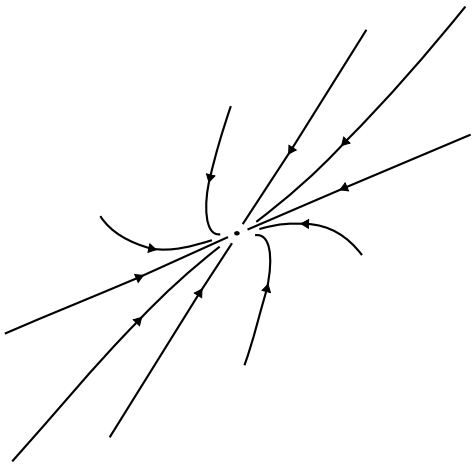


Figura 1.3: Un nodo estable.

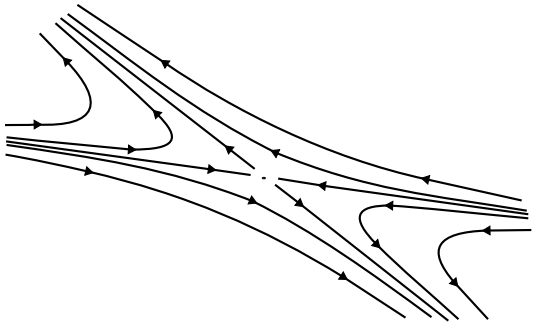


Figura 1.4: Un punto silla.

miento aparenta estar organizado por el punto crítico, por esta razón a estos comportamientos en torno a puntos críticos en dos dimensiones se les dan los nombres de nodo inestable, nodo estable y punto silla respectivamente, aplicándose estos nombres tanto al comportamiento de las trayectorias, como a los puntos críticos que presentan en su entorno dichos comportamientos.

En los casos en que los dos eigenvalores sean complejos estos tendrán que ser complejos conjugados y no habrá un sistema de coordenadas en el que la matriz sea diagonal, sin embargo las soluciones podrán expresarse como combinaciones lineales de las exponenciales de los eigenvalores por el tiempo. Expresando los eigenvalores como $\lambda_1 = r + i\omega$ y $\lambda_2 = r - i\omega$ obtendremos

$$\begin{aligned} x_1 &= e^{rt}(c_1 e^{i\omega t} + c_2 e^{-i\omega t}), \\ x_2 &= e^{rt}(c_3 e^{i\omega t} + c_4 e^{-i\omega t)}, \end{aligned} \tag{1.30}$$

naturalmente habrá que ajustar los coeficientes complejos c para obtener soluciones reales.

Como podemos ver si la parte real de los eigenvalores r es positiva las soluciones se alejarán exponencialmente del origen o punto crítico y si es negativa las soluciones se acercarán asintóticamente a dicho punto. Por otro lado la combinación de las exponenciales imaginarias conjugadas nos darán una rotación en torno del origen, por lo que al combinarse ambos comportamientos obtendremos que las trayectorias son espirales en torno al punto crítico, las que tenderán asintóticamente a él si r es negativo o al infinito si es positivo, este tipo de soluciones se ilustra en la figura 1.5

A estos comportamientos generales, así como a los puntos críticos que presentan trayectorias similares a éstas en su entorno, se les llama focos, estable al primero e inestable al segundo. Notemos que, tanto para los focos como para los nodos, el calificativo de estable e inestable corresponde con las ideas de equilibrio estable e inestable cuando vemos al punto crítico como una situación de equilibrio, ya que en el caso estable al comenzar en un punto cercano al crítico nos acercamos exponencialmente a él, mientras que en el inestable nos alejamos.

Cuando la parte real es cero tendremos únicamente la parte de las exponenciales imaginarias conjugadas y por tanto las trayectorias serán curvas cerradas concéntricas con el origen. Naturalmente toda trayectoria cerrada en el espacio fase representa soluciones periódicas de la ecuación diferencial, el oscilador armónico es un ejemplo de este caso. A este comportamiento así como a los puntos críticos que lo presentan en su entorno se les llama centro y en general se ven como se ilustra en la figura 1.6

De esta forma hemos logrado clasificar las ecuaciones diferenciales lineales en dos dimensiones al encontrar todos sus posibles comportamientos

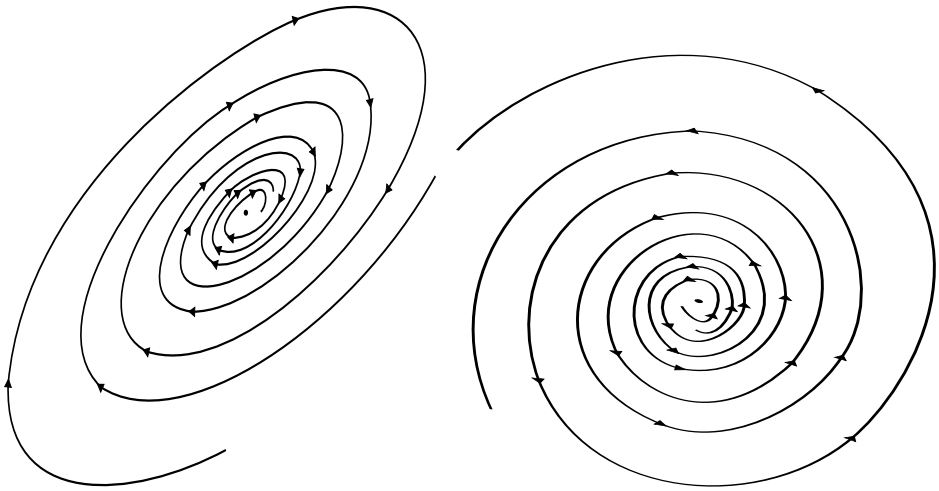


Figura 1.5: Un foco estable y uno inestable.

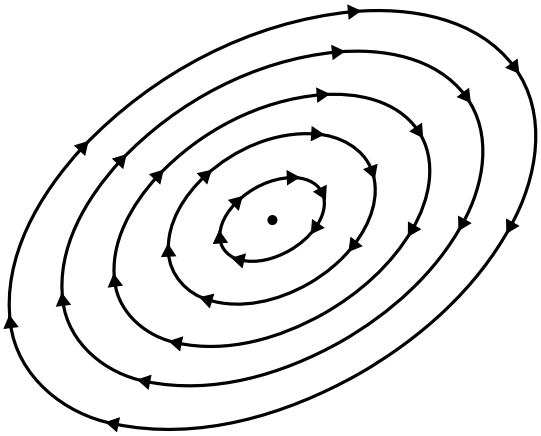


Figura 1.6: Un centro.

generales o topológicos y hemos mostrado esquemas que ilustran esos comportamientos en el espacio fase. A estos dibujos los llamamos en general diagramas fase y son sumamente útiles para ilustrar el comportamiento de las soluciones de las ecuaciones diferenciales. Naturalmente hay algunos casos sumamente excepcionales que no hemos tratado, aquellas ecuaciones en que la matriz tiene dos eigenvalores iguales, el lector podrá sin dificultad extender este estudio a esos casos excepcionales.

Para más de dos dimensiones el estudio anterior puede ser generalizado sin mayores dificultades. El origen será también un punto crítico y será el único que exista, pero, el número de eigenvalores de la matriz que define a la ecuación lineal crecerá con la dimensión y en consecuencia crecerá el número de posibles combinaciones entre los que son reales positivos o negativos y los complejos conjugados con parte real positiva, negativa o cero.

En el espacio fase se presentará una serie de planos y rectas que pasan por el origen y que se intersectan transversalmente, correspondientes, los primeros, a las parejas de eigenvalores complejos conjugados y las segundas a los eigenvalores reales. Sobre cada plano las trayectorias permanecerán en él, o sea serán planos invariantes en el sentido de que si las condiciones iniciales se ubican en un punto del espacio fase sobre uno de esos planos, todo el desarrollo de la trayectoria quedará sobre él, además será una espiral, ya sea estable o inestable o, en el caso en que la parte real de los eigenvalores sea cero, una curva cerrada concéntrica con el origen. Cada una de las semirrectas correspondientes a eigenvalores reales será, por si misma, una trayectoria asintótica al origen si el eigenvalor es negativo o tendiente al infinito si es positivo. Las demás trayectorias, o sea las correspondientes a condiciones iniciales generales que no se ubican en los casos anteriores, serán curvas que combinen los comportamientos ya descritos. Como ejemplo en la figura 1.7 hemos puesto una trayectoria en tres dimensiones para el caso en que se tenga un eigenvalor real positivo y un par de complejos conjugados con parte real negativa. La extensión a los casos excepcionales con eigenvalores iguales puede hacerse con facilidad.

De esta forma hemos podido hacer un estudio general de las ecuaciones diferenciales lineales, las hemos clasificado, esto es hemos encontrado todos los posibles comportamientos y hemos construido sus diagramas fase. Esto nos permite tener una visión geométrica o topológica de sus soluciones o trayectorias.

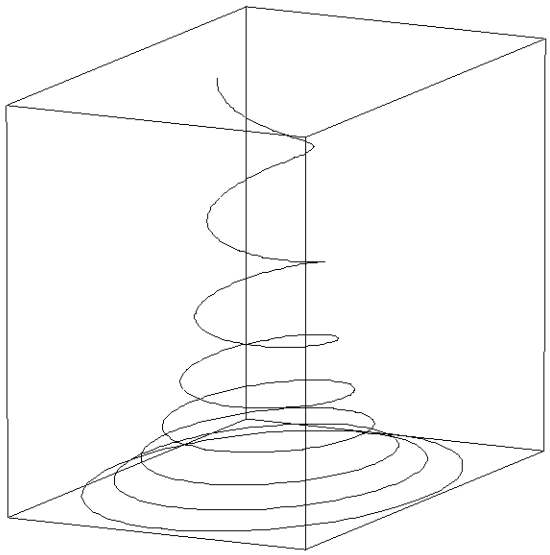


Figura 1.7: Ejemplo de una trayectoria de ecuaciones lineales en tres dimensiones.

1.3. Ecuaciones diferenciales no lineales.

En el caso de las ecuaciones diferenciales nolineales no existen métodos sistemáticos para resolverlas ni para clasificarlas o encontrar sus comportamientos cualitativos o topológicos generales. Para ver con mayor claridad la complejidad que introduce la nolinealidad veamos algunas cuestiones básicas al respecto.

Tomemos una ecuación diferencial

$$\dot{x} = f(x) \tag{1.31}$$

donde $x \in R^n$ y $f : R^n \to R^n$. Un punto crítico de esta ecuación es un punto p del espacio fase en el que se cumple la ecuación; para que esto ocurra se requiere que $f(p) = 0$, esto es, los puntos críticos de una ecuación diferencial son las raíces de la ecuación algebraica $f(x) = 0$ o, lo que es lo mismo, los puntos en los que el campo vectorial que define a la ecuación se hace cero.

Para las ecuaciones lineales vimos que siempre existe uno de estos puntos, el origen, y únicamente puede existir uno. Para el caso no lineal puede no existir ninguno o pueden existir varios puntos críticos, de hecho para encontrarlos debemos encontrar las raíces de un sistema de ecuaciones algebraicas no lineales y sabemos que no hay métodos sistemáticos para hacerlo, ni tan sólo los hay para saber cuantas raíces tiene uno de estos sistemas; he aquí una primera diferencia importante entre las ecuaciones lineales y las no lineales, en el caso de las lineales el origen y sólo el origen es punto crítico, en el caso no lineal no tenemos métodos sistemáticos para determinar los puntos críticos. Claro está, dada una ecuación particular, en muchas ocasiones podemos encontrar los puntos críticos, pero en general no lo podemos hacer, y como conocer los puntos críticos es un primer paso en el conocimiento del comportamiento de las soluciones de una ecuación diferencial, esta primera dificultad nos hace ver ya la complejidad a la que conduce la no linealidad.

Conocer los puntos críticos de una ecuación diferencial constituye un gran avance en su estudio ya que nos permite acercarnos al comportamiento de sus soluciones aplicando el siguiente procedimiento:

Digamos que p es un punto crítico de la ecuación $\dot{x} = f(x)$ donde $x \in R^n$ y $f : R^n \to R^n$, o sea, $f(p) = 0$. Esto nos permite, usando el teorema de Taylor, expresar f en una vecindad de p como

$$\begin{aligned} f_1(p+\varepsilon) &= \left.\frac{\partial f_1}{\partial x_1}\right|_p \varepsilon_1 + \left.\frac{\partial f_1}{\partial x_2}\right|_p \varepsilon_2 + \cdots + \left.\frac{\partial f_1}{\partial x_n}\right|_p \varepsilon_n + O(|\varepsilon|^2), \quad (1.32) \\ &\vdots \\ f_n(p+\varepsilon) &= \left.\frac{\partial f_n}{\partial x_1}\right|_p \varepsilon_1 + \left.\frac{\partial f_n}{\partial x_2}\right|_p \varepsilon_2 + \cdots + \left.\frac{\partial f_n}{\partial x_n}\right|_p \varepsilon_n + O(|\varepsilon|^2), \end{aligned}$$

o lo que es lo mismo

$$f(p+\varepsilon) = Df_p\varepsilon + O(|\varepsilon|^2) \quad (1.33)$$

donde Df_p significa la diferencial de f valuada en p, que es la transformación lineal representada por la matriz Jacobiana de $f(x)$ valuada en p .

Si $\mid \varepsilon \mid \ll 1$, esto es dentro de una vecindad suficientemente pequeña del punto crítico, podemos despreciar los términos de orden superior en la ecuación (1.33) y podremos aproximar la ecuación diferencial por una ecuación lineal[3]. Esto nos enseña que en una vecindad suficientemente pequeña de un punto crítico, las trayectorias o soluciones de la ecuación diferencial se verán como las de una ecuación lineal definida por la diferencial de f valuada en ese punto crítico y presentará alguno de los comportamientos generales estudiados anteriormente.

De esta forma podemos conocer el comportamiento de las soluciones de las ecuaciones diferenciales en torno a vecindades suficientemente pequeñas de los puntos críticos. Presentamos en la figura 1.8 un ejemplo hipotético

Naturalmente esto únicamente nos permite estudiar la ecuación en torno a los puntos críticos, pero sabemos que el resto del espacio fase debe estar lleno de trayectorias que no se pueden intersectar y que deben unirse o extenderse de manera continua y congruente con los comportamientos lineales ya establecidos en las vecindades de los puntos críticos.

Por otro lado dada una configuración de puntos críticos pueden existir muchas formas topológicamente distintas de llenar el espacio con trayectorias que cumplan con lo señalado anteriormente, como un ejemplo de esto en la figura 1.9 se muestran tres de las formas en que esto se puede hacer para el caso ilustrado en que tenemos dos puntos silla.

Es importante señalar que en el caso de dos dimensiones las trayectorias, que son curvas, dividen al espacio fase, mientras que en más de dos dimensiones, al ser las trayectorias también curvas, ya no dividen el espacio, lo que

[3]Estamos suponiendo que no ocurre alguna situación excepcional con la matriz Jacobiana del sistema, en cuyo caso la parte lineal no es suficiente.

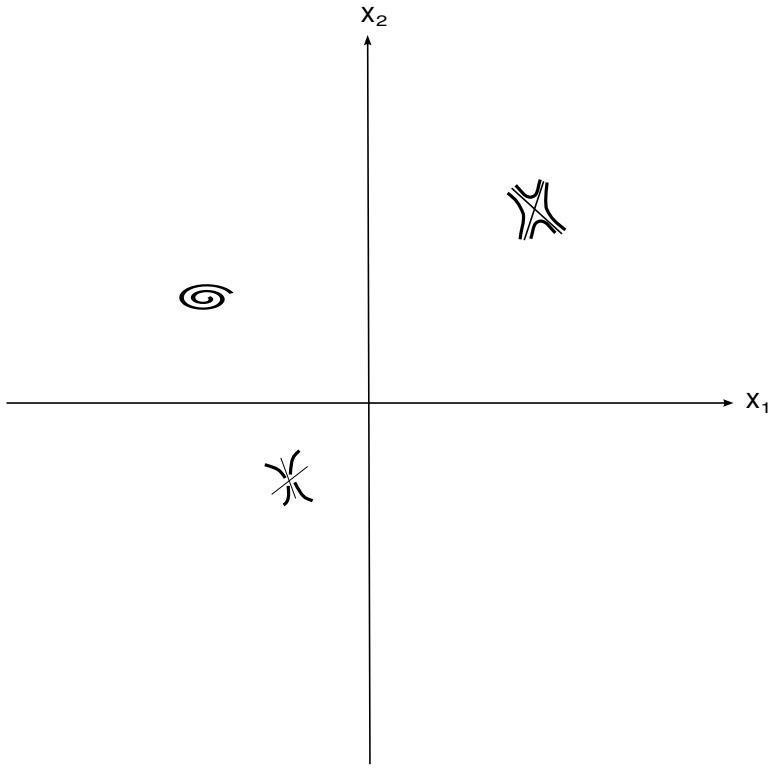


Figura 1.8: Una ecuación que presenta tres puntos críticos, un nodo, un punto silla y un foco.

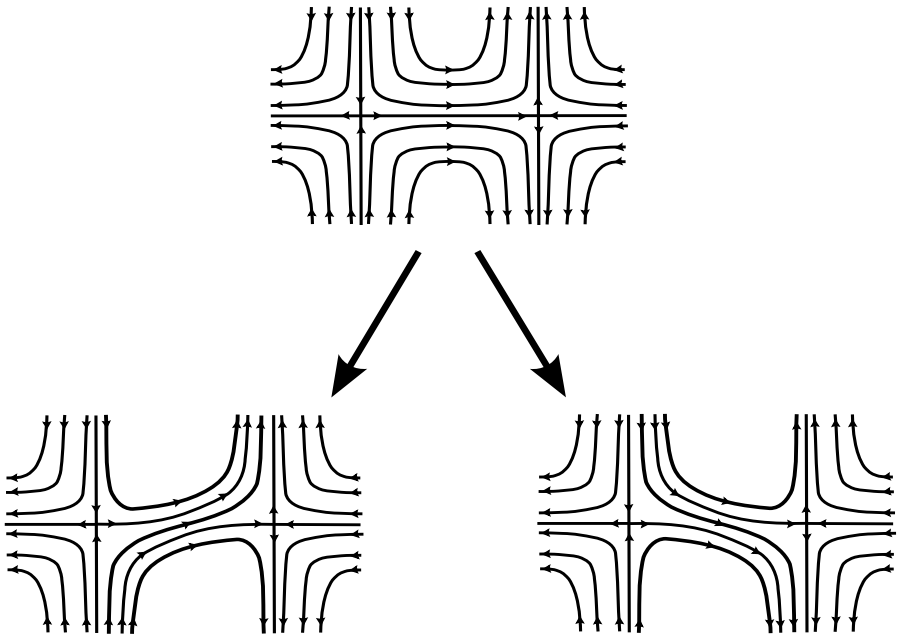


Figura 1.9: Diversas formas de unir, en el espacio fase, dos puntos silla.

da libertades adicionales, aumentando considerablemente la complejidad e incluso abriendo posibilidades de comportamientos que no existen en dos dimensiones y que hacen del estudio de las ecuaciones diferenciales no lineales un campo abierto a la investigación.

1.4. El concepto de flujo en el espacio fase.

Las soluciones o trayectorias de una ecuación diferencial son curvas en el espacio fase parametrizadas por el tiempo y, como hemos dicho, por cada punto de éste pasa alguna de ellas, esto nos permite hacer la siguiente construcción:

Tomemos un valor del tiempo digamos cero y un cierto punto del espacio fase x, después de un tiempo t el punto x, al recorrer la trayectoria que pasa por él, estará en otro punto del espacio fase, digamos x'. Como esto pasa para cada punto del espacio fase, dado un valor del tiempo t podemos establecer un mapeo del espacio fase en si mismo

$$\psi_t : R^n \to R^n \tag{1.34}$$

definido por

$$\psi_t : x \longmapsto x', \tag{1.35}$$

esto es, a cada punto del espacio fase se le asocia aquel punto al que irá a parar cuando recorre su trayectoria por un tiempo t. Generalizando esto podemos construir un mapeo $\psi : R^n \times R \to R^n$ que a cada punto del espacio fase $x \in R^n$ y a cada tiempo $t \in R$, les asocia el punto del espacio fase $x' \in R^n$ según lo establecido por el mapeo ψ_t.

Estos mapeos, ya sea el conjunto de todos los ψ_t o el que los comprende a todos ψ, representan las soluciones de la ecuación diferencial, ya que si tenemos estos mapeos, tenemos el conjunto de soluciones y viceversa, si tenemos las soluciones se pueden construir los mapeos, por lo que son equivalentes a la ecuación diferencial y podemos decir que el estudio de este tipo de mapeos es otra forma de ver a las ecuaciones diferenciales. Por su semejanza a la descripción del fluir de un liquido a estos mapeos se les llama el flujo en el espacio fase y presentan la ventaja de poder contemplar en un sólo objeto matemático todas las soluciones de la ecuación diferencial.

El conjunto de mapeos ψ_t tiene propiedades interesantes, veamos:

En primer lugar es claro que ψ_0 es el mapeo idéntico, ya que el punto al que irá a parar cualquier punto después de transcurrido un tiempo cero es él mismo. En segundo lugar la composición de dos de estos mapeos

$$\psi_{t_1} \circ \psi_{t_2} \tag{1.36}$$

será un tercero de ellos y es fácil ver que se trata de $\psi_{t_1+t_2}$. En tercer lugar dado uno de ellos ψ_t, existe otro, el correspondiente a recorrer la trayectoria en el sentido negativo del tiempo, ψ_{-t}, tal que al componerlo con el original da el mapeo idéntico. Finalmente, al igual que para cualquier conjunto de mapeos, la operación de composición es asociativa. Concluimos por tanto que el conjunto de mapeos ψ_t forma grupo ante la operación de composición; es más, notamos que la estructura de este grupo es la misma que la estructura del grupo de los números reales ante la adicióno, puesto que la composición de dos de los mapeos corresponde al mapeo determinado por un valor del tiempo que es la suma de los tiempos que se asocian a los dos mapeos originales y al elemento idéntico de los reales, el cero, se le asocia el elemento idéntico del conjunto de mapeos; decimos entonces que el flujo es un grupo de un parámetro de mapeos del espacio fase en si mismo.

Esta propiedad del flujo es, desde luego, consecuencia de haber sido construido a partir de las soluciones de las ecuaciones diferenciales, pero lo inverso es también cierto. Esto es, si tengo un conjunto de mapeos de R^n en R^n parametrizados por los números reales tales que, ante la composición, forman

un grupo cuya estructura es la del grupo de los reales ante la suma, entonces conforman un flujo y habrá alguna ecuación diferencial en R^n cuyas soluciones o trayectorias generen ese flujo o, lo que es lo mismo, cuyas trayectorias sean las lineas de flujo correspondientes al grupo de mapeos. A un conjunto de mapeos parametrizados de la forma descrita se le llama también una acción del grupo de los reales sobre el espacio R^n.

Así vemos que las ecuaciones diferenciales son equivalentes a las acciones del grupo de los reales en R^n, por lo que el estudio de estas acciones es equivalente al estudio de las ecuaciones diferenciales, o al de los flujos o al de los campos vectoriales y podemos siempre aprovechar las ventajas conceptuales y técnicas que cada una de estas formas de contemplar las ecuaciones diferenciales nos brindan.

Hemos visto en este capítulo un breve panorama de la teoría de las ecuaciones diferenciales ordinarias, debemos advertir que nuestra intención ha sido hacer una descripción para introducir algunos conceptos que serán de utilidad en el estudio de la mecánica clásica. En ningún momento hemos tenido la intención de ser precisos o formales, ni mucho menos de abarcar el basto campo de las ecuaciones diferenciales; al lector interesado en ampliar esta exposición lo remitimos a la bibliografía especializada.

Capítulo 2

La mecánica vectorial.

La mecánica y el cálculo diferencial fueron formulados originalmente por Newton en términos geométricos siguiendo las formas utilizadas en el trabajo matemático de su época. En el contexto moderno la formulación física y matemáticamente más cercana a la forma newtoniana es la que conocemos como formulación vectorial, la que se presenta en los cursos elementales.

Newton establece, en primer término, la presencia del espacio físico y del tiempo como sustrato físico del universo y, por tanto, de su teoría del movimiento de la materia. Sustenta su mecánica en un espacio tridimensional, continuo, euclidiano y absoluto, que constituye el sistema de referencia primario al que toda medición debe finalmente reducirse. Sobre el tiempo piensa de manera similar, lo contempla como un transcurrir continuo, inmutable y absoluto que a todo afecta de manera simultanea y que es medible por el correr uniforme de un parámetro universal. Este espacio y este tiempo newtonianos forman parte integral de la teoría y, si bien, se llega a hacer uso de diversos espacios matemáticos a lo largo del desarrollo de la mecánica, ese espacio tridimensional y ese tiempo de la mecánica newtoniana son los únicos que tiene realidad física y a ellos hay que reducir, a fin de cuentas, toda referencia.

La materia para la mecánica de Newton es reducible siempre al concepto de partícula, esto es, al de un objeto puntual cuyas únicas propiedades físicas son: estar sujeta inexorablemente al transcurrir del tiempo físico, ocupar en cada instante de ese tiempo un lugar preciso y medible en el espacio y tener una masa medible. Desde la perspectiva de la mecánica clásica, esto es de la teoría newtoniana del movimiento de la materia en el espacio, todo cuerpo material se reduce a un conjunto de partículas entre las que se establecen relaciones e interacciones. Estos son los sustentos básicos de la mecánica clá-

sica y, por ello, la mecánica es también una teoría de la materia, el espacio y el tiempo. Estas ideas inciden en el estudio de todo lo que llamamos física y llegan a establecerse como base de nuestra concepción general de la naturaleza y el universo.

En otro nivel de la teoría mecánica aparecen diversos conceptos, unos derivados o referidos de estos, como lo son el de velocidad, aceleración y sistema inercial de referencia, otros derivables del concepto primario de fuerza como medida de la interacción entre partículas.

Todo ello se conjuga para integrar una teoría del movimiento por medio de las tres leyes de Newton. La primera establece la propiedad primaria de la inercia galileana; la segunda descubre relaciones de causa y efecto entre las propiedades espaciales y temporales de las partículas, la masa y las fuerzas; la tercera, condiciona la interacción entre partículas a una reciprocidad total,lo que evita distingo alguno entre ellas más allá de su propia calidad definitoria, la masa.

Estos son los conceptos físicos básicos que estructuran toda la mecánica clásica, pueden ser expresados de muy diversas formas, por ejemplo substituyendo algo de lo aquí dicho por una ley de conservación o puede todo esto estructurarse matemáticamente de muy diversas formas, pero el conjunto conceptual físico, siguirá siendo el mismo. Podemos decir que las dos grandes 'revoluciones' de la física del siglo XX son revoluciones por el cambio profundo que producen en este esquema conceptual. La relatividad por modificar los conceptos de espacio y tiempo, integrándolos, correlacionándolos y liberándolos de su carácter absoluto; la mecánica cuántica dejando atrás el concepto de materia reducible a partículas definidas únicamente por su masa y posición en el espacio.

La formulación vectorial de la mecánica refleja de manera directa los conceptos newtonianos. El espacio se contempla como un espacio vectorial de tres dimensiones con métrica euclidiana, el tiempo como un parámetro real, la masa como un escalar y las fuerzas como vectores. No nos detendremos aquí en el desarrollo de esta formulación de la mecánica, la damos por estudiada como un antecedente de este texto. Nuestro principal objetivo es el desarrollo de las formulaciones llamadas analíticas y el estudio de las propiedades del movimiento a que nos conduce la teoría mecánica. Sin embargo, en este capítulo, analizaremos, utilizando la formulación vectorial, algunos sistemas mecánicos importantes; nuestro objetivo es fijar ideas, notaciones y revisar los métodos de la mecánica vectorial con el fin de recordar sus ventajas y descubrir sus puntos débiles, puntos que, en gran parte, inspiran la busqueda de otras formulaciones de la teoría física del movimiento de la materia en el espacio y el tiempo.

2.1. Sistemas mecánicos en una dimensión.

Los sistemas mecánicos más simples son los constituidos por una partícula de masa m que se encuetra sobre una recta, que está sometida a una fuerza dirigida sobre esa misma recta y cuya magnitud sólo depende de la posición de la propia partícula.

En este caso la segunda ley de Newton nos dice que

$$m\frac{d^2x}{dt^2} = F(x), \tag{2.1}$$

donde x denota la posición de la partícula sobre la recta y $F(x)$ es la magnitud de la fuerza que actúa sobre ella; supondremos que la función F es continua y derivable.

Esta ecuación, como todas las de la mecánica clásica, es de segundo orden, pero podemos construir un espacio fase para ella y ponerla en la forma de un sistema de ecuaciones de primer orden. Para esto usamos como variable adicional a la velocidad de la partícula, esto es $v = dx/dt$, y, siguiendo lo hecho en el capítulo anterior, obtenemos

$$\begin{aligned} \frac{dx}{dt} &= v \\ \frac{dv}{dt} &= \frac{1}{m}F(x). \end{aligned} \tag{2.2}$$

Debido a que estamos en una dimensión, siempre será posible encontrar un potencial, esto es una función $V(x)$ tal que

$$F(x) = -\frac{\partial V(x)}{\partial x}. \tag{2.3}$$

Esto implica la existencia de una cantidad conservada, en este caso una integral de movimiento, la energía total, que nos permitirá resolver de manera general esta clase de sistemas mecánicos.

Se define la energía total como la suma de dos componentes: la energía cinética y la potencial. La energía potencial es la función $V(x)$ y la energía cinética, que denotaremos por T, se define por medio de

$$T = \frac{1}{2}mv^2. \tag{2.4}$$

Así la energía total E estará dada por

$$E = \frac{1}{2}mv^2 + V(x). \tag{2.5}$$

Como decíamos está cantidad física es una constante de movimiento, es decir una cantidad que depende de las variables dinámicas y del tiempo tal que a lo largo de la evolución temporal o movimiento del sistema permanece constante. En general dada una ecuación diferencial en cierto espacio fase M, diremos que una función $f : M \times R \to R$ del espacio fase y el tiempo en los reales es una constante de movimiento si es continua, derivable, está bien definida en todo M y permanece constante a lo largo de las trayectorias o soluciones de la ecuación diferencial; naturalmente ese valor constante dependerá de la trayectoria o, lo que es equivalente, de las condiciones iniciales. En el importante caso en que esa función no dependa explícitamente del tiempo la llamaremos integral de movimiento.

Demostremos que la energía total es una integral de movimiento. Calculemos para hacerlo su derivada total con respecto al tiempo. Notamos que E no depende explícitamente del tiempo, pero lo hace por medio de x y v que si dependerán del tiempo al recorrerse la trayectoria del sistema mecánico, entonces

$$\frac{dE}{dt} = \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial v}\frac{dv}{dt} = \frac{\partial V}{\partial x}\frac{dx}{dt} + mv\frac{dv}{dt}, \tag{2.6}$$

pero las derivadas de x y de v con respecto al tiempo a lo largo de las trayectorias están dadas por la ecuación de movimiento

$$\begin{aligned} \frac{dx}{dt} &= v \\ \frac{dv}{dt} &= -\frac{1}{m}\frac{dV}{dx}, \end{aligned} \tag{2.7}$$

y al sustituirlas en la ecuación (2.6) obtenemos $dE/dt = 0$, que es lo que queríamos demostrar.

Dadas unas condiciones iniciales para el movimiento, x_0 y v_0, la energía tomará un valor E_0 que a lo largo del movimiento permanecerá constante. Esto nos permite encontrar directamente la curva solución en el espacio fase usando la ecuación

$$\frac{1}{2}mv^2 + V(x) = E_0. \tag{2.8}$$

Así, para cada valor de la energía total tendremos una curva, podemos, entonces, usar la energía para fijar una de las dos condiciones iniciales, esto determina la trayectoria. La otra condición inicial determinará el punto sobre la trayectoria, o fase, en la que se encuentra el sistema en el tiempo que hayamos escogido como tiempo inicial. Naturalmente esto último tiene una importancia física menor.

Si deseamos encontrar explícitamente la trayectoria, despejamos v de la ecuación y recordando que $v = dx/dt$ obtenemos

$$\frac{dx}{dt} = \sqrt{\frac{2(E_0 - V(x))}{m}} \tag{2.9}$$

e integrando[1]

$$t = \int_{x_0}^{x} \sqrt{\frac{m}{2(E_0 - V(x))}} dx. \tag{2.10}$$

Esto nos proporciona $x(t)$ y derivando obtenemos $v(t)$. Vemos que todo sistema mecánico en una dimensión se puede reducir a cuadraturas, es decir, al cálculo explícito de algunas integrales.

En muchas ocasiones resulta interesante y útil tener una descripción cualitativa del sistema mecánico en su conjunto, esto se puede hacer mediante la obtención de un diagrama fase, esto es, el trazado de un conjunto de trayectorias en el espacio fase que nos permite visualizar el comportamiento general del sistema. En el caso de los sistemas mecánicos en una dimensión podemos utilizar un procedimiento gráfico para construir ese diagrama fase. Pasemos a describirlo.

Tomamos un sistema coordenado en el que trazamos la gráfica de la función potencial $V(x)$. Debajo de ésta gráfica colocamos un sistema de coordenadas para el espacio fase, de tal manera que concuerden los ejes de las ordenadas. (Ver figura 2.1)

Consideremos un conjunto de valores para la energía total $E_1, E_2, \cdots, E_n$, que, como hemos establecido, permanecerán constantes a lo largo de las trayectorias. Estos valores de la energía quedan representados por rectas punteadas sobre la gráfica de la energía potencial. En cada valor de x la diferencia entre el valor de la energía total y la energía potencial será la energía cinética y en la gráfica será la distancia vertical entre la curva que representa al potencial y la recta correspondiente al valor de la energía total. Naturalmente esta diferencia no puede ser negativa, ya que esto implicaría valores imaginarios de la velocidad, por lo que solo habrá movimiento en las regiones en que la curva potencial esté por debajo de la recta de la energía total.

Cuando el potencial presente un mínimo y la energía total sea igual al potencial evaluado en ese mínimo, la energía cinética, y por tanto la velocidad, serán cero; cuando el sistema esté en esa situaciónten permanecerá estático o en equilibrio sobre el punto donde el potencial alcanza valor mínimo, esto es, con velocidad cero. En la representación gráfica del espacio fase toda la

[1] Hemos tomado el tiempo inicial t_0 como cero, cuestión que siempre haremos a no ser que se indique otra cosa.

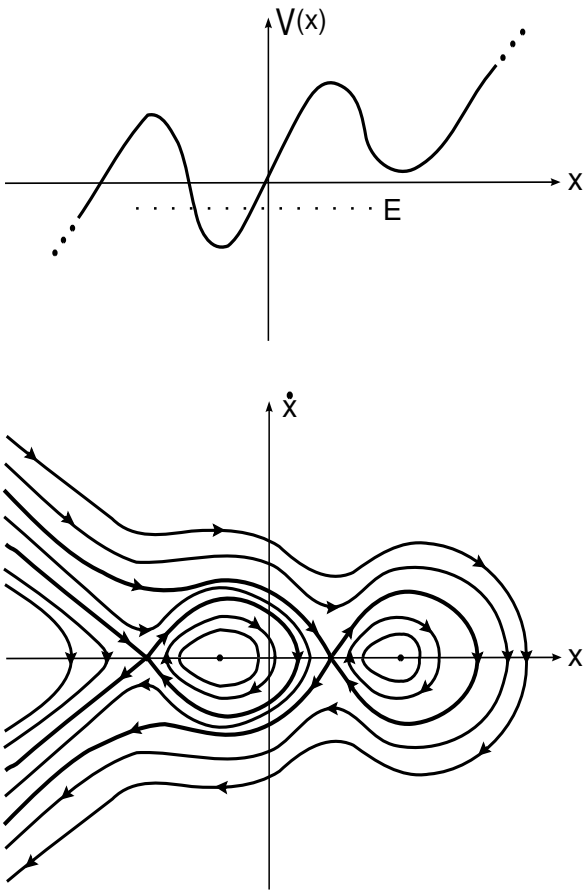


Figura 2.1: En el método gráfico para trazar las trayectorias se coloca la gráfica del potencial en la parte superior y la del espacio fase en la inferior. Los máximos y mínimos del potencial se corresponden con puntos críticos en el espacio fase y las trayectorias se obtienen a partir de la conservación de la energía total.

trayectoria será un punto sobre el eje x colocado justo debajo del mínimo del potencial, esto es, a cada mínimo del potencial le corresponde un punto crítico en el espacio fase.

Para una energía total un poco superior a este mínimo, existirá un intervalo de valores de x en los que puede estar el sistema y como en la energía cinética tenemos el cuadrado de la velocidad, para cada valor de x dentro de ese intervalo podrá haber dos valores de la velocidad, uno correspondiente a la raíz positiva y el otro a la negativa; la trayectoria en el espacio fase será entonces una curva cerrada, simétrica con respecto al eje x, que lo intersecta verticalmente y que, con el transcurrir del tiempo, se recorre en el sentido de las manecillas del reloj.

Así, en todo un entorno del punto correspondiente al mínimo, tendremos en el espacio fase una serie de trayectorias cerradas y concéntricas, cada una de ellas corresponde a un movimiento oscilatorio y periódico entre dos valores extremos de x, aquellos en que la energía potencial iguala a la total. En el lenguaje del capítulo anterior diríamos que la presencia de un mínimo en el potencial implica la existencia de un punto crítico en el espacio fase, este punto crítico es un centro.

En caso de tener un máximo del potencial, el sistema puede permanecer en la posición del máximo con velocidad cero si la energía total es igual a la potencial en ese máximo, esto es el sistema permanecerá estático o en equilibrio y estará representado por una trayectoria consistente en un solo punto del espacio fase de manera similar a lo que ocurre en un mínimo. Sin embargo esa es toda la similitud, ya que ahora, debido a que estamos en un máximo, en un entorno de esa posición y para la misma energía total, la diferencia entre la energía total y la potencial será positiva y podrá haber movimiento; esto se verá en el espacio fase con la existencia de cuatro trayectorias asintóticas al punto de equilibrio, dos que tienden hacia él y dos que se le alejan, naturalmente habrá simetría con respecto al eje x. Al aumentar o disminuir un poco la energía total notamos la ocurrencia de dos situaciones, pues al aumentarla tendremos soluciones que se recorren sin cambio en el sentido de la velocidad, mientras que al disminuirla se forma una "barrera de potencial" y se generan trayectorias en las que se da un cambio en el sentido de la velocidad. Al trazar estas trayectorias en el espacio fase vemos que se forma un punto silla en torno al equilibrio. Es importante señalar que el punto que representa la situación de equilibrio es de por si una trayectoria, no pasan por él dos trayectorias que se cortan, son, en realidad cinco trayectorias distintas que no tienen puntos en común.

Toda función continua y derivable en una dimensión presenta una secuencia de máximos y mínimos (salvo la ocurrencia de puntos de inflexión,

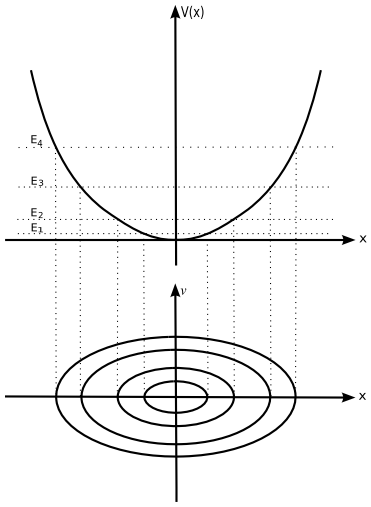


Figura 2.2: Método gráfico para obtener las trayectorias del oscilador armónico.

que por ser estructuralmente inestables y no presentar mayores dificultades, no trataremos aquí) y un cierto comportamiento al infinito, así conocer las trayectorias en torno a los puntos extremales de la función potencial y las alturas relativas entre los máximos y mínimos, nos permite extender las trayectorias en el espacio fase por medio de la construcción gráfica que estamos haciendo y obtener el diagrama fase completo para el sistema mecánico en estudio. En la figura 2.1 se muestra un ejemplo hipotético que ilustra este proceso y muestra los comportamientos generales que podemos esperar.

2.1.1. El oscilador armónico.

Veamos, como ejemplo, el caso del oscilador armónico. La energía potencial en este caso es

$$V(x) = \frac{1}{2}kx^2, \tag{2.11}$$

o sea un parábola cuyo único extremo es un mínimo colocado en el origen. En la figura 2.2 presentamos la construcción general del diagrama fase para este sistema.

Como podemos ver todas las trayectorias son curvas cerradas concéntricas que corresponden a soluciones periódicas.

Las ecuaciones de movimiento para este caso son

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -\frac{k}{m}x, \end{aligned} \tag{2.12}$$

que es una ecuación lineal en dos dimensiones, cuyo campo vectorial fue ilustrado en la figura 1.1; vemos que efectivamente $x = 0$ y $v = 0$ es un punto crítico y que los eigenvalores de la matriz que define a la ecuación son $\pm i\sqrt{k/m}$, por lo que deducimos que se trata de un centro, lo que concuerda con lo obtenido en el análisis gráfico.

La energía total es

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2, \tag{2.13}$$

la ecuación (2.10), que reduce a cuadraturas todo sistema unidimensional, en este caso es

$$t = \int_{x_0}^{x} \sqrt{\frac{m}{2E - kx^2}}dx = \sqrt{\frac{m}{k}} \arcsin\left(\sqrt{\frac{k}{2E}}x\right) - \sqrt{\frac{m}{k}}\beta, \tag{2.14}$$

donde $\beta = \arcsin\left(\sqrt{k/2E}x_0\right)$, de aquí encontramos que

$$x = \sqrt{\frac{2E}{k}} \sin\left(\sqrt{\frac{k}{m}}t + \beta\right) \tag{2.15}$$

y derivando

$$v = \sqrt{\frac{2E}{m}} \cos\left(\sqrt{\frac{k}{m}}t + \beta\right), \tag{2.16}$$

la conocida solución del oscilador armónico. Naturalmente al graficar en el espacio las trayectorias dadas por estas ecuaciones, obtenemos elipses concéntricas en concordancia con lo obtenido en el método gráfico.

2.1.2. El péndulo simple.

Veamos otro ejemplo, el del péndulo simple. Una varilla indeformable y sin masa que es capaz de girar en torno a uno de sus extremos, quedando en el otro extremo una masa m actuada por la gravedad, como se muestra en la figura 2.3.

En términos de la coordenada θ la función potencial es $V(\theta) = -mgh\cos\theta$ que presenta mínimos en los valores $\theta_l = \pm\pi l$ para l cero o par y máximos para l impar. El análisis gráfico será el que se muestra en la figura 2.4 .

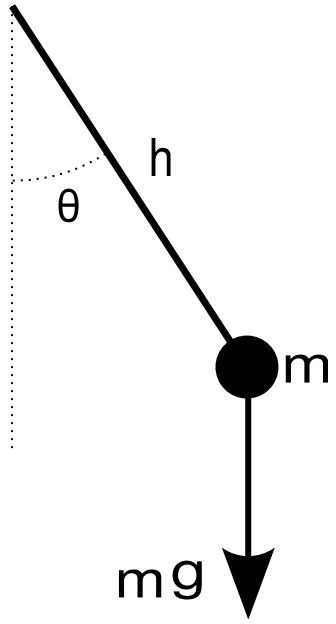


Figura 2.3: El péndulo simple.

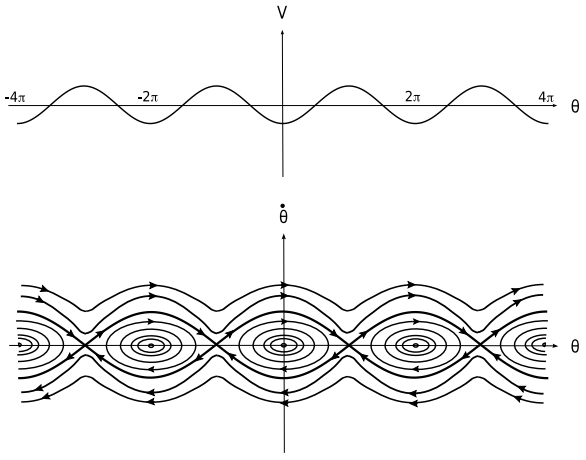


Figura 2.4:

Sabemos que el péndulo tiene dos puntos de equilibrio, la posición inferior y la superior, sin embargo en el espacio fase vemos una infinidad de ellos. Esto se debe a que hemos representado la posición como si fuera una recta o un espacio R cuando en la realidad este espacio es un círculo que no es topológicamente equivalente a R. Si el espacio que representa la situación geométrica del sistema mecánico no es topológicamente equivalente a R, entonces el espacio fase no lo será a R^2, ya que éste es el producto cartesiano del espacio geométrico por el que representa a la velocidad, que en general si es R. En el caso del oscilador armónico, tanto el espacio geométrico, como el de velocidades, eran R y el espacio fase $R \times R = R^2$ por lo que no tuvimos problemas. En el caso del péndulo, el espacio geométrico es el círculo S^1, y el de velocidades R, por tanto el espacio fase será $S^1 \times R$ que es un cilindro y no R^2.

Ahora bien el proceso gráfico que hemos utilizado, puede hacerse sin mayores dificultades sobre un cilindro. Tomemos un cilindro, su circunferencia representa al círculo que constituye la geometría del péndulo, esto es, el conjunto de posibles posiciones de la partícula, y la posición a lo largo de su eje representa la velocidad de la partícula, este cilindro es el espacio fase. Arriba de este cilindro ponemos otro en el que trazamos la gráfica del potencial en función de la posición del péndulo, obtenemos una elipse que viene a ser la intersección del cilindro con un plano inclinado. Vemos que al recorrer la circunferencia del cilindro nos encontramos con un solo mínimo y un solo máximo diametralmente opuestos. Las rectas que indicaban el valor de la energía total serán ahora círculos que le dan vuelta al cilindro y podremos hacer en la parte inferior, que representa al espacio fase, exactamente la misma construcción que en R^2. La figura 2.5 muestra claramente esta construcción.

Como podemos ver el punto que representa la posición inferior y estática del péndulo es un punto crítico rodeado de curvas cerradas concéntricas que representan las oscilaciones periódicas del péndulo. A medida que se aumenta la energía total, aumenta la amplitud de estas curvas y por tanto la de la oscilación. Para energías suficientemente grandes el círculo que representa a la energía total, queda colocado arriba de la elipse que es la gráfica de la función potencial, no hay puntos donde la energía total y la potencial sean iguales y por tanto donde la energía cinética se haga cero, esto implica que el movimiento se da en torno del cilindro; en el espacio fase vemos curvas cerradas que se cierran al dar una vuelta al cilindro, unas en la dirección positiva en la parte superior, otras en la dirección negativa en la parte inferior. Estas curvas cerradas son también soluciones periódicas, pero corresponden al movimiento en vuelta de campana y no a una oscilación. Finalmente te-

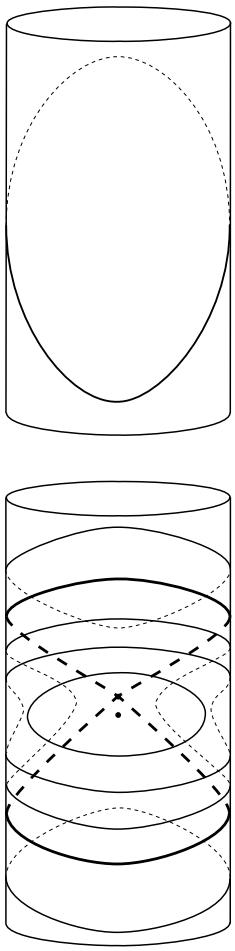


Figura 2.5: El espacio fase de un péndulo es un cilindro. En la figura superior se gráfica el potencial $V(\theta) = -mgh\cos\theta$ y en la inferior se trazan en el espacio fase algunas de las trayectorias del péndulo. Las oscilaciones en torno al punto de equilibrio inferior aparecen como curvas cerradas colapsables en un punto sin salirse del cilindro; las trayectorias correspondientes a vueltas de campana son curvas cerradas que le dan una vuelta al cilindro y no son colapsables en un punto. Con un trazo más grueso se han destacado las separatrices que aparentan intersectarse en el punto de equilibrio superior, éste queda, desde el punto de vista de la figura, en la parte posterior del cilindro.

nemos el caso en que la energía total tiene el mismo valor que el máximo de la energía potencial; aquí tenemos una trayectoria que es la del péndulo en equilibrio en la posición vertical, ésta queda representada por un punto crítico con la configuración de un punto silla; las dos trayectorias que salen de él, dan la vuelta al cilindro y regresan al punto crítico, claro está, salen y llegan al punto crítico únicamente de manera asintótica. Estas trayectorias muestran claramente que un péndulo al que se le imprime una energía total igual a la máxima energía potencial, realiza un movimiento en el que siempre ira subiendo hacia el punto de equilibrio superior, pero que lo hace con velocidad decreciente, de tal manera que nunca llega a él.

Como en cualquier espacio fase las trayectorias cerradas representan soluciones periódicas, pero vemos que en este caso tenemos dos tipos distintos de curvas cerradas, las que representan oscilaciones y las que representan vueltas de campana. La diferencia física entre estos movimientos se ve en el espacio fase como una diferencia topológica, las curvas cerradas correspondientes a oscilaciones periódicas pueden contraerse a un punto sin salirse del espacio, esto es del cilindro, mientras que las otras no; de hecho sobre el cilindro todas las que representan oscilaciones periódicas son deformables unas en las otras, lo mismo podemos decir de las que representan vueltas de campana, pero unas no son deformables en las otras. Diferencias debidas a cuestiones topológicas en el espacio fase, como éstas, se presentan en muchas ocasiones, no sólo en una dimensión, sino en general. Llamaremos libraciones a los movimientos periódicos que se representan en el espacio fase por curvas cerradas que se pueden deformar hasta reducirlas a un punto y rotaciones a aquellos cuya trayectoria en el espacio fase no puede ser deformada hasta llevarla a un punto. Esta nomenclatura proviene de la mecánica celeste.

Notemos también que las trayectorias correspondientes a una energía total igual a un máximo del potencial, dividen las regiones del espacio fase donde se dan ambos tipos de movimiento. Por esta razón se les llama separatrices.

Poco más se puede decir sobre el movimiento mecánico en una dimensión. Debido a que todo espacio unidimensional es topológicamente equivalente a una recta, esto es a R, o a un círculo S^1, el espacio fase lo será a un plano R^2 , o a un cilindro $S^1 \times R$ y entonces una figura como la 2.1, o su equivalente sobre un cilindro, resume todos los comportamientos posibles del movimiento en una dimensión.

2.2. El movimiento en dos o más dimensiones.

Al pasar a sistemas en dos o más dimensiones ya no es posible, por medio de un procedimiento sistemático, reducir a cuadraturas todos los sistemas mecánicos como ocurrió en el caso de una dimensión. Analicemos el caso general de una partícula de masa m que se mueve sobre la superficie de un plano.

Denotemos el vector de posición de la partícula con respecto a un sistema de referencia inercial arbitrario por $\mathbf{r}$[2] y a la fuerza que actúa sobre ella por $\mathbf{F}(\mathbf{r},t)$, entonces de la segunda ley de Newton obtenemos la ecuación de movimiento

$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r},t). \tag{2.17}$$

En un sistema cartesiano de coordenadas (x,y) esta ecuación se expresará de la siguiente manera

$$m\ddot{x} = F_x(x,y,t) \tag{2.18}$$
$$m\ddot{y} = F_y(x,y,t), \tag{2.19}$$

y llamando $v_x = \dot{x}$ y $v_y = \dot{y}$ obtenemos un espacio fase de cuatro dimensiones (x,y,v_x,v_y). En este espacio la ecuación de movimiento será

$$\dot{x} = v_x \tag{2.20}$$
$$\dot{y} = v_y \tag{2.21}$$
$$\dot{v}_x = \frac{1}{m}F_x \tag{2.22}$$
$$\dot{v}_y = \frac{1}{m}F_y. \tag{2.23}$$

En una dimensión encontramos que siempre es posible derivar la fuerza de una función potencial, lo que condujo a la construcción de una integral de movimiento, la energía total. En el caso de dos dimensiones, diremos que se ha definido una función potencial cuando podamos expresar la fuerza, ahora con su carácter vectorial, como menos el gradiente de dicha función. Esto es cuando

$$\mathbf{F}(\mathbf{r},t) = -\mathrm{grad}V(\mathbf{r},t), \tag{2.24}$$

[2]Usaremos letras negritas para denotar vectores. Esto lo haremos únicamente en este capítulo, en el resto del texto se indicará en cada caso la notación para los vectores. Esto se hace así para destacar el papel principal que tienen los vectores en la formulación vectorial.

o, en coordenadas cartesianas,

$$F_x(x,y,t) = -\frac{\partial V(x,y,t)}{\partial x} \tag{2.25}$$

$$F_y(x,y,t) = -\frac{\partial V(x,y,t)}{\partial y}. \tag{2.26}$$

En una dimensión bastó que la fuerza fuera una función continua para que existiera el potencial, en el caso de dos dimensiones para que un campo vectorial sea derivable de un potencial se requiere en adición que

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}. \tag{2.27}$$

Encontramos así una diferencia importante entre los sistemas mecánicos en una dimensión y en dos o más: en estos últimos no necesariamente existe una función potencial.

En el caso en que exista un potencial podemos definir la energía total por medio de

$$E = \frac{1}{2}m(\mathbf{v}\cdot\mathbf{v}) + V(\mathbf{r},t), \tag{2.28}$$

y, en el caso en que la función potencial no dependa explícitamente del tiempo, esta energía total será una integral de movimiento. La demostración de esto es similar en dos o más dimensiones a la hecha para el caso de una dimensión. En adición a esto, la presencia de una dimensión más, aún en los casos en que la energía total esté definida y sea una integral de movimiento, impide la reducción general del sistema a cuadraturas. Esto constituye una diferencia fundamental entre los sistemas en varias dimensiones y los de una sola dimensión.

Estas cuestiones, y el hecho de que en general las ecuaciones diferenciales que nos permiten estudiar los sistemas mecánicos no son lineales, nos hacen ver que, en más de una dimensión, no existen procedimientos sistemáticos para resolver los problemas mecánicos. Esto nos obliga a estudiar algunas clases de problemas solubles, a estudiar casos concretos importantes por procedimientos específicos o a buscar desarrollos teóricos que nos permitan profundizar en las cuestiones conceptuales y eventualmente encontrar métodos más poderosos de solución.

Pasemos pues a estudiar algunos ejemplos ilustrativos e importantes de sistemas en dos o más dimensiones.

2.2.1. Figuras de Lissajous.

Un primer ejemplo, muy simple, consiste en una partícula de masa m capaz de moverse en un plano y sujeta a una fuerza de atracción dirigida hacia una recta y proporcional a su distancia a dicha recta, simultáneamente es atraída hacia otra recta, perpendicular a la primera, con una magnitud proporcional a la distancia que la separa de esa segunda recta. La fuerza resultante es la adición vectorial de estas dos, por lo que su dirección y no únicamente su magnitud depende de la posición de la partícula. Poniendo un sistema de coordenadas cartesianas cuyos ejes coincidan con dichas rectas, la fuerza será

$$\mathbf{F} = -k_x x\hat{i} - k_y y\hat{j}, \tag{2.29}$$

donde $\hat{i}$ y $\hat{j}$ son los vectores unitarios que forman la base natural para coordenadas cartesianas. La ecuación de movimiento será

$$m\mathbf{a} = \mathbf{F}, \tag{2.30}$$

la que, en coordenadas cartesianas, se expresa en componentes como

$$\begin{aligned} m\ddot{x} &= -k_x x \\ m\ddot{y} &= -k_y y. \end{aligned}$$

Estas ecuaciones resultan desacopladas por lo que su solución es simplemente la superposición de las soluciones de cada una de ellas, esto es, dos osciladores armónicos. Podemos ver que la curva que la partícula traza en el plano (x, y) son las conocidas figuras de Lissajous. El espacio fase tiene cuatro dimensiones (x, y, v_x, v_y), las figuras de Lissajous no son otra cosa que las proyecciones en el plano (x, y) de las trayectorias en ese espacio fase. Las soluciones serán periódicas si las dos frecuencias $\omega_x = \sqrt{k_x/m}$ y $\omega_y = \sqrt{k_y/m}$ son conmensurables, esto es, si existen enteros n_x y n_y distintos de cero tales que $\omega_x n_x + \omega_y n_y = 0$. En caso contrario, esto es, si las frecuencias son inconmensurables, la solución no será periódica; en este caso al completarse un ciclo de una de las variables, digamos x, la otra avanza siempre por una misma proporción irracional, por esta razón se les llama soluciones cuasiperiódicas.

En términos de las trayectorias, tanto en el espacio fase como en el físico, las periódicas corresponden a curvas cerradas y las cuasiperiódicas a curvas abiertas; estas últimas, como ocurre con las figuras de Lissajous, llenan densamente cierta región del espacio; esto es, dado cualquier punto de

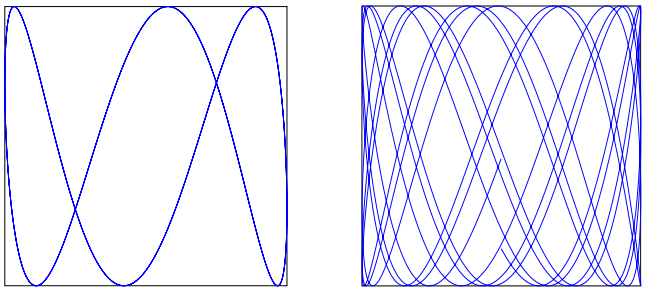


Figura 2.6: Figuras de Lissajous. A la izquierda un caso conmensurable o periódico, a la derecha uno inconmensurable o cuasiperiódico.

dicha región y una vecindad arbitraria de él, la curva pasará necesariamente por esa vecindad, mas no necesariamente por el punto. En este ejemplo las frecuencias están determinadas por los parámetros físicos del sistema, la masa de la partícula y las constantes de los resortes, por lo tanto el que se tengan soluciones periódicas o cuasiperiódicas no depende de las condiciones iniciales; esto es, para un sistema dado todas las soluciones serán periódicas o todas serán cuasiperiódicas.

En este ejemplo tenemos un sistema conservativo, esto es, existe una función potencial y la energía total es una integral de movimiento. Obviamente la función potencial es $V(x,y) = \frac{1}{2}k_x x^2 + \frac{1}{2}k_y y^2$. En este caso adicionalmente a la energía, e independientemente de ella, podemos tener otra integral de movimiento, por ejemplo $\frac{1}{2}mv_x^2 + \frac{1}{2}k_x x^2$.

2.2.2. El campo central.

Un segundo ejemplo en dos dimensiones es el de una partícula de masa m sometida a una fuerza central, esto es, a una fuerza que siempre está en la dirección que une a la partícula con un punto fijo del espacio y cuya magnitud es función únicamente de la distancia entre la partícula y ese punto fijo o centro de fuerzas. Este ejemplo es de gran importancia tanto por su contenido físico como histórico.

Puede decirse que éste es un sistema mecánico en tres dimensiones y no en dos, pero dadas unas condiciones iniciales se reduce a un sistema en dos dimensiones. Esto puede argumentarse de la siguiente manera: Tomemos un punto del espacio tridimensional y fijemos en él el centro de fuerzas, pongamos la partícula en algún otro punto del espacio y démosle una velocidad vectorial inicial; esto define el sistema mecánico y da las seis condiciones iniciales que determinan la trayectoria de la partícula. El plano que contie-

ne al centro de fuerzas, a la partícula y al vector velocidad inicial contiene también al vector fuerza, por tanto no puede haber una componente de la aceleración perpendicular a ese plano, como la componente de la velocidad inicial perpendicular al plano es cero, la componente perpendicular al plano de la velocidad de la partícula será siempre cero, o sea la partícula se moverá siempre sobre ese plano. Otra forma de decir lo mismo consiste en señalar que: debido a que el vector fuerza siempre pasa por el centro de fuerza, la torca con respecto a ese centro es cero y por tanto el vector impulso angular es constante; ahora el plano que contiene a la partícula, al centro de fuerza y al vector velocidad siempre es perpendicular al vector de impulso angular, por lo que ese plano es invariante. Por ello estudiaremos este sistema como uno en dos dimensiones sin perder generalidad.

La fuerza central tendrá, en coordenadas cartesianas, las componentes

$$F_x = \frac{x}{r}F(r) \tag{2.31}$$

y

$$F_y = \frac{y}{r}F(r), \tag{2.32}$$

donde $r = \sqrt{x^2+y^2}$ es la distancia al centro de fuerzas y $F(r)$ la magnitud de la fuerza. Podemos ver fácilmente que cumple con 2.27 y por tanto derivable de un potencial $V\left(\sqrt{x^2+y^2}\right)$.

Naturalmente la simetría ante rotaciones de la fuerza central nos indica que en coordenadas polares (r,θ) la solución será más simple. En estas coordenadas el potencial se expresará simplemente como $V(r)$.

Las ecuaciones, como siempre, deben expresarse en forma vectorial, por lo que, para ponerlas en términos de componentes, debemos tomar en cuenta que en los sistemas de coordenadas no cartesianos los vectores que forman la base cambian al desplazarnos de un punto del espacio a otro, esto se ilustra en la figura 2.7 en que comparamos la situación entre un sistema cartesiano y uno polar.

La ecuación vectorial de movimiento en este caso es

$$m\ddot{\mathbf{r}} = \mathbf{F}(r), \tag{2.33}$$

donde $\mathbf{r} = r\hat{\mathbf{r}}$ y $\mathbf{F}(r) = F(r)\hat{\mathbf{r}}$ y hemos usado $\hat{\mathbf{r}}$ y $\hat{\theta}$ para denotar a los vectores unitarios que forman la base asociada a las coordenadas polares. Ahora debemos calcular $\ddot{\mathbf{r}}$, para esto primero calcularemos $\dot{\mathbf{r}}$ a partir de que $\mathbf{r} = r\hat{\mathbf{r}}$, entonces

$$\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt}, \tag{2.34}$$

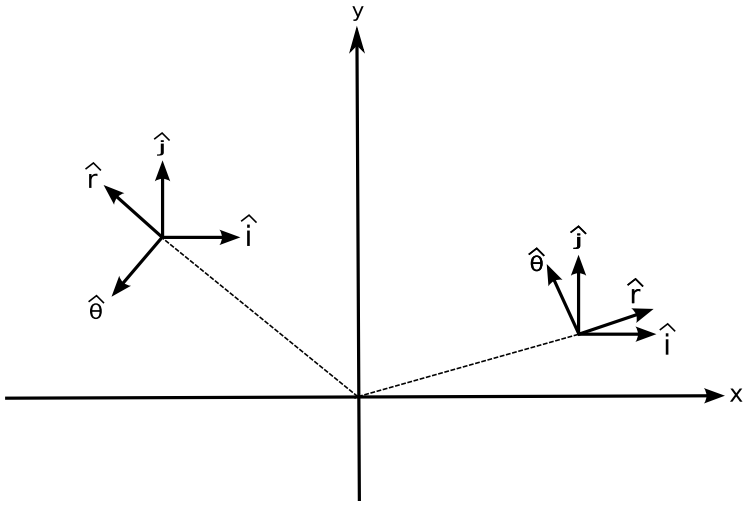


Figura 2.7: Los vectores de las bases inducidas por las coordenadas cartesianas y las polares en R^2.

por lo que debemos calcular $d\hat{\mathbf{r}}/dt$. De la figura 2.8 se ve claramente que $d\hat{\mathbf{r}}/dt = \dot{\theta}\hat{\theta}$, por lo que

$$\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\theta}. \tag{2.35}$$

Ahora tendremos que

$$\ddot{\mathbf{r}} = \ddot{r}\hat{\mathbf{r}} + \dot{r}\frac{d\hat{\mathbf{r}}}{dt} + \dot{r}\dot{\theta}\hat{\theta} + r\ddot{\theta}\hat{\theta} + r\dot{\theta}\frac{d\hat{\theta}}{dt}, \tag{2.36}$$

y de la misma figura 2.8 encontramos que $d\hat{\theta}/dt = -\dot{\theta}\hat{\mathbf{r}}$, por lo que

$$\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\mathbf{r}} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\theta}, \tag{2.37}$$

y como la fuerza esta dada por $\mathbf{F} = F(r)\hat{\mathbf{r}}$ entonces la ecuación de movimiento 2.33 expresada en componentes polares será

$$m(\ddot{r} - r\dot{\theta}^2) = F(r) = -\frac{\partial V(r)}{\partial r}, \tag{2.38}$$

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = 0. \tag{2.39}$$

La segunda de estas ecuaciones, la que corresponde a la componente angular, puede expresarse como

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0, \tag{2.40}$$

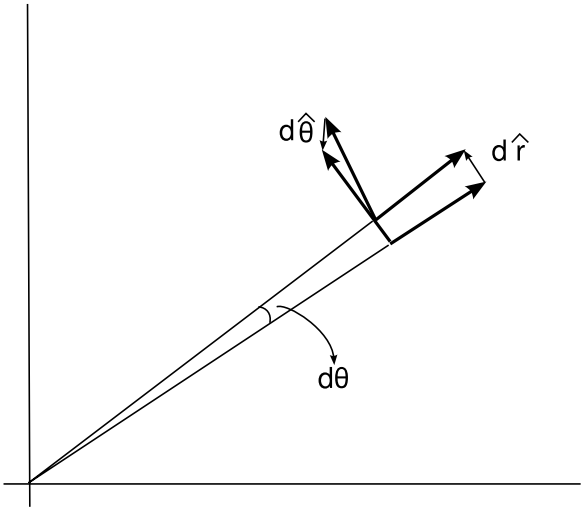


Figura 2.8: La diferencial de los vectores de la base en coordenadas polares.

lo que nos hace ver que, en adición a la energía total, existe otra integral de movimiento o cantidad conservada; es inmediato identificarla con la magnitud del impulso angular total $l = mr^2\dot{\theta}$. A lo largo de las trayectorias del sistema l permanece constante en el valor determinado por las condiciones iniciales.

El haber encontrado una integral de movimiento adicional nos permite simplificar el problema de resolver las ecuaciones de movimiento, ya que podemos despejar $\dot{\theta}$ en función de l y r, esto es

$$\dot{\theta} = \frac{l}{mr^2} \tag{2.41}$$

y podemos substituir este valor en la ecuación de movimiento correspondiente a la componente radial para obtener

$$m\ddot{r} = \frac{l^2}{mr^3} + F(r) = -\frac{d}{dr}\left(\frac{l^2}{2mr^2} + V(r)\right). \tag{2.42}$$

Esta ecuación es fácilmente soluble ya que multiplicando ambos lados por $\dot{r}$ obtenemos

$$m\ddot{r}\dot{r} = -\frac{d}{dr}\left(\frac{l^2}{2mr^2} + V(r)\right)\frac{dr}{dt}, \tag{2.43}$$

de aquí podemos ver que

$$\frac{d}{dt}\frac{1}{2}m\dot{r}^2 = -\frac{d}{dt}\left(\frac{l^2}{2mr^2} + V(r)\right) \tag{2.44}$$

esto es

$$\frac{1}{2}m\dot{r}^2 = -\frac{l^2}{2mr^2} - V(r) + E \tag{2.45}$$

donde E es una constante de integración, que obviamente es la energía total.

Naturalmente esta última ecuación corresponde a la conservación de la energía total y hubiésemos podido escribirla directamente, pues es valida para todo sistema conservativo. Podemos ahora despejar y obtener

$$\frac{dr}{dt} = \sqrt{\frac{2}{m}\left[E - V(r) - \frac{l^2}{2mr^2}\right]}, \tag{2.46}$$

al integrar obtenemos que

$$\int_{r_0}^{r} \frac{dr}{\sqrt{E - V(r) - \frac{l^2}{2mr^2}}} = \sqrt{\frac{2}{m}}t, \tag{2.47}$$

donde r_0 es el valor de r al tiempo inicial $t = 0$. Esto nos proporciona r como una función del tiempo y finalmente usando (2.41) obtenemos

$$\theta = \theta_0 + \int_0^t \frac{l}{mr(t)^2}dt, \tag{2.48}$$

donde θ_0 es el valor inicial de θ. Hemos reducido así las ecuaciones de movimiento a cuadraturas, es decir al cálculo de integrales explícitas. Las cuatro condiciones iniciales necesarias para determinar la solución se expresan en las constantes E, l, r_0 y θ_0.

Es interesante ver que la ecuación (2.42) es totalmente equivalente a la ecuación de movimiento de una partícula de masa m sujeta a una fuerza derivada de un potencial de la forma

$$V_{ef}(r) = \frac{l^2}{2mr^2} + V(r), \tag{2.49}$$

razón por la que lo llamamos potencial efectivo[3]. La reducción a este sistema unidimensional nos permite usar las técnicas gráficas desarrolladas en el capítulo anterior para dilucidar el comportamiento radial del movimiento. Por otro lado la ecuación (2.41) nos hace ver que la coordenada angular siempre

[3]En general se utiliza esta nomenclatura cuando se reduce un sistema mecánico a otro equivalente de menor dimensionalidad.

crece o decrece, en dependencia del signo de l[4] que es invariante, por lo que al conjugar el comportamiento de r y el de θ podemos construir una imagen cualitativa de las órbitas sin tener que realizar las integrales en las ecuaciones 2.47 y 2.48.

Por ejemplo si el potencial $V(r)$ es tal que el potencial efectivo $V_{ef}(r)$ presenta un mínimo para un cierto valor r_0, entonces para una energía total cercana al valor de ese mínimo tendremos, como se ilustra en la figura 2.9, que r oscilará entre un valor $r_{\text{mín}}$ y un valor $r_{\text{máx}}$; simultáneamente θ variará monótonamente, por lo que la órbita será la combinación de una rotación en torno al centro de fuerza y una oscilación radial periódica que acerca y aleja a la partícula de dicho centro entre un distancia mínima y una máxima; el resultado será una roseta como la indicada en la figura 2.10.

Por razones históricas llamamos a la distancia mínima el pericentro y a la máxima apocentro, podemos ver que la velocidad angular de la partícula será mínima en el apocentro y máxima en el pericentro. Cuando el ángulo que subtienden dos pericentros (o apocentros) consecutivos es conmensurable con 2π la órbita será cerrada y el movimiento periódico; cuando dicho ángulo no sea conmensurable con 2π el movimiento será cuasiperiódico, la órbita no se cerrará y cubrirá densamente la región comprendida entre los dos círculos; para una energía total dada el que dicho ángulo sea conmensurable con 2π o no lo sea dependerá del valor de l, como lo indica la ecuación (2.48). Si la energía total coincide con el valor del potencial efectivo en un mínimo de este, la órbita será circular.

Cuando el potencial efectivo no acota al movimiento radial la órbita se va al infinito y hay que hacer un cálculo explícito usando la ecuación (2.48) para determinar si, en su carrera hacia el infinito, la partícula realiza alguna o algunas revoluciones completas en torno al centro de fuerza o no alcanza a hacerlo.

Un caso de gran importancia es el de la fuerza gravitacional o sea el potencial $V(r) \sim r^{-1}$ y negativo. La gráfica del potencial efectivo se muestra en la figura 2.11

Podemos ver que para energías totales negativas el movimiento será acotado, o sea la órbita será una roseta, para energías positivas el movimiento será abierto.

[4]El caso en que $l = 0$ la órbita está sobre una recta que pasa por el centro de fuerzas y el sistema se reduce a una dimensión. En adición a esto puede ocurrir una colisión con dicho centro, el que, como pasa en nuestro caso, puede presentar una singularidad del potencial. Para el proposito de nuestra revisión de la mecánica vectorial la reducción de estas posibles órbitas singulares no es de importancia, por lo que supondremos que en todo caso $l \neq 0$.

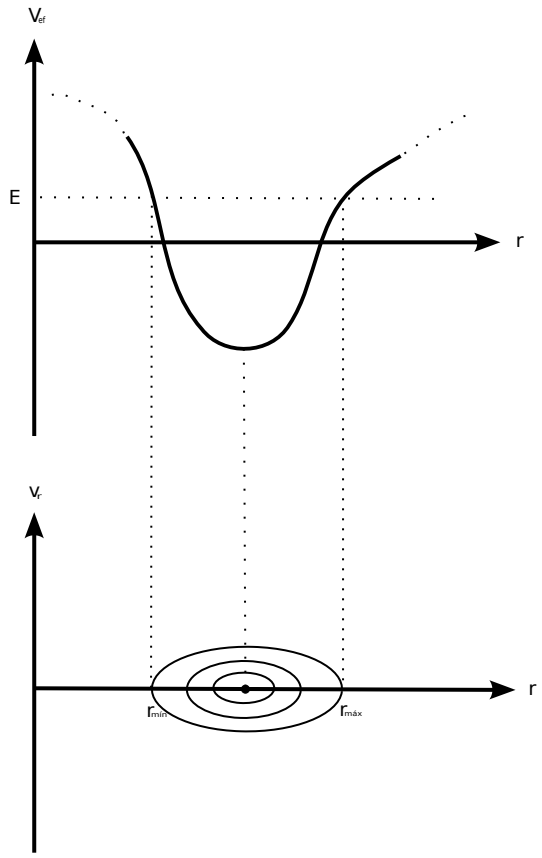


Figura 2.9: Comportamiento de r en un entorno de un mínimo del potencial efectivo para un campo central.

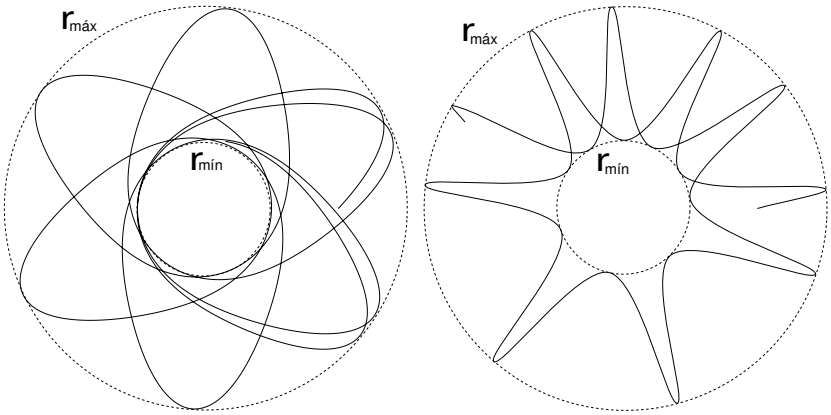


Figura 2.10: En un potencial central para energías cercanas a un mínimo del potencial efectivo las órbitas son rosetas. La razón entre las frecuencias de los movimientos radial y angular determina la forma de estas rosetas.

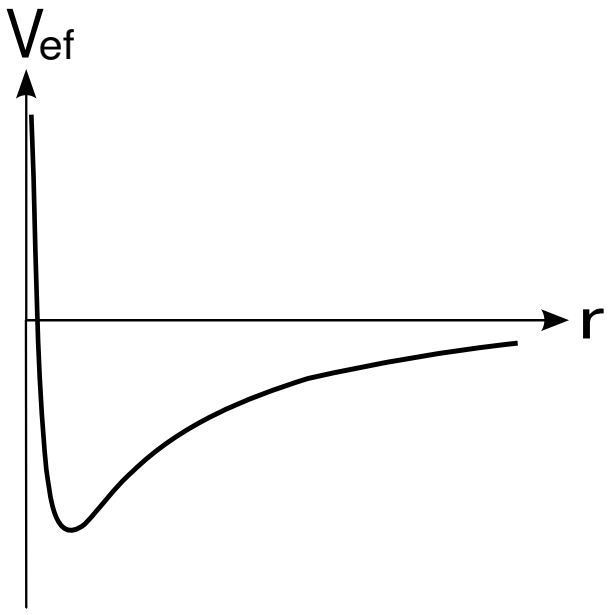


Figura 2.11: El potencial efectivo correspondiente a un potencial gravitacional.

Veamos con detalle este caso.

Obtendremos en primer lugar, para el caso general de un potencial central, una ecuación diferencial para la órbita, esto implica eliminar la variable t entre las ecuaciones (2.38) y ver a r como una función de θ. Tendremos que

$$\frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt}, \tag{2.50}$$

pero por (2.41)

$$\frac{d\theta}{dt} = \frac{l}{mr^2} \tag{2.51}$$

lo que nos da

$$\frac{dr}{dt} = \frac{l}{mr^2}\frac{dr}{d\theta} \tag{2.52}$$

y aplicando el procedimiento otra vez

$$\frac{d^2r}{dt^2} = \frac{l^2}{m^2r^2}\frac{d}{d\theta}\left(\frac{1}{r^2}\frac{dr}{d\theta}\right). \tag{2.53}$$

Usando (2.42) la ecuación de la órbita será

$$\frac{l^2}{mr^2}\frac{d}{d\theta}\left(\frac{1}{r^2}\frac{dr}{d\theta}\right) = \frac{l^2}{mr^3} + F(r), \tag{2.54}$$

la que al tomar en consideración que

$$\frac{1}{r^2}\frac{dr}{d\theta} = -\frac{d}{d\theta}\left(\frac{1}{r}\right) \tag{2.55}$$

y llamando u a $1/r$ nos da

$$\frac{l^2u^2}{m}\left(\frac{d^2u}{d\theta^2} + u\right) = -F(1/u). \tag{2.56}$$

Para el caso kepleriano $F(1/u) = -ku^2$ y la ecuación 2.56 será la ecuación lineal

$$\frac{d^2u}{d\theta^2} + u - \frac{mk}{l^2} = 0, \tag{2.57}$$

cuya solución es

$$\frac{1}{r} = u = \frac{mk}{l^2} + c\cos(\theta - \varphi), \tag{2.58}$$

donde c y φ son constantes de integración que quedarán determinadas por las condiciones iniciales. Está es la ecuación de la órbita en coordenadas polares y podemos observar que definiendo ε por

$$\varepsilon = \frac{l^2}{mk}c, \tag{2.59}$$

se puede poner en la forma

$$\frac{1}{r} = \frac{mk}{l^2}\left(1 + \varepsilon\cos(\theta - \varphi)\right), \tag{2.60}$$

que, en coordenadas polares, es la ecuación general de una cónica de excentricidad ε con un foco en el origen, esto es, en el centro de fuerzas.

Podemos expresar la excentricidad de la órbita en función de la energía total y del impulso angular. Para ello recordamos de 2.45 que, cuando $\dot{r} = 0$,

$$E = \frac{1}{2}\frac{l^2}{mr^2} - \frac{k}{r}, \tag{2.61}$$

pero, cuando $\dot{r} = 0$, $\theta = \varphi$ y el valor de r será

$$\frac{1}{r_\varphi} = \frac{mk}{l^2}(1 + \varepsilon), \tag{2.62}$$

tendremos entonces que

$$E = \frac{mk^2}{2l^2}(1 + \varepsilon)^2 - \frac{mk^2}{l^2}(1 + \varepsilon), \tag{2.63}$$

de donde encontramos que

$$1 + \varepsilon = 1 \pm \sqrt{1 + 2\frac{l^2}{mk^2}E}, \tag{2.64}$$

o, como la excentricidad debe ser siempre positiva,

$$\varepsilon = \sqrt{1 + 2\frac{l^2}{mk^2}E}. \tag{2.65}$$

Vemos que para energías negativas la excentricidad será menor que uno y las órbitas serán elipses y para energías positivas la excentricidad será mayor que uno y las órbitas hipérbolas; el caso intermedio, el de energía cero, nos da excentricidad uno y órbitas parabólicas. Naturalmente la energía no podrá ser menor que

$$-\frac{mk^2}{2l^2} \tag{2.66}$$

pues de serlo la excentricidad sería imaginaria, para el valor de la energía correspondiente a este mínimo la excentricidad es cero y la órbita será circular; este valor es, desde luego, el valor mínimo del potencial efectivo dado en (2.49). Podemos ver que cuando las órbitas son acotadas siempre son elipses, esto es, rosetas de una sola hoja y por tanto la órbita siempre es periódica, independientemente de las condiciones iniciales. Es interesante señalar que sólo existen dos potenciales centrales para los cuales todas las órbitas acotadas son periódicas, el gravitacional y el potencial de la forma kr^2 con k positiva. El hecho de que que para el caso planetario las órbitas son elipses uno de cuyos focos está ocupado por el Sol se conoce como la primera Ley de Kepler.

Pasemos a calcular el período τ de las órbitas elípticas. De la ecuación (2.51) obtenemos que

$$dt = \frac{m}{l} r^2 d\theta, \tag{2.67}$$

ahora $r^2 d\theta$ es dos veces el área que subtiende el radiovector $\mathbf{r}$ al recorrerse el ángulo por $d\theta$, por tanto

$$dt = 2\frac{m}{l} dA. \tag{2.68}$$

Notamos que esta ecuación implica que el radiovector $\mathbf{r}$ recorre áreas iguales en tiempos iguales y que esta propiedad es válida para toda fuerza central ya que la hemos deducido a partir únicamente de la conservación del impulso angular. Esta relación se conoce, para el caso planetario, como la segunda ley de Kepler. En el caso del potencial $-k/r$ esta ecuación nos hace ver que el período de la órbita elíptica es proporcional al área de la elipse ya que al transcurrir un período el radiovector barre toda la elipse; como el área de una elipse está dada por πab, donde a y b son los semiejes mayor y menor de la elipse, obtenemos

$$\tau = 2\pi \frac{m}{l} ab. \tag{2.69}$$

El semieje mayor de la elipse es la mitad de la suma de los valores máximo y mínimo de r, que pueden obtenerse de la ecuación 2.60, encontramos que

$$\sqrt{1-\varepsilon^2} = \frac{l}{\sqrt{mk}} \frac{1}{\sqrt{a}} \tag{2.70}$$

Ahora, la relación entre los semiejes de una elipse está dada por

$$b = a\sqrt{1-\varepsilon^2}, \tag{2.71}$$

lo que nos permite, usando (2.69), calcular el período

$$\tau = 2\pi \sqrt{\frac{m}{k}} a^{\frac{3}{2}}. \tag{2.72}$$

En el caso planetario $k = GMm$ donde G es la constante de gravitación universal, M la masa del Sol y m la del planeta, por lo que, para un planeta

$$\tau = 2\pi\sqrt{\frac{1}{GM}}a^{\frac{3}{2}}, \tag{2.73}$$

que es la tercera ley de Kepler para el movimiento planetario, el período no depende de la masa del planeta y varia en proporción directa al cubo e inversa al cuadrado del semieje mayor de la órbita.

En el caso kepleriano en adición al vector de impulso angular se presenta otro vector que permanece constante frente al movimiento: el vector de Runge Lenz. Este vector se suele denotar con la letra $\mathbf{A}$ y esta definido por

$$\mathbf{A} = m\dot{\mathbf{r}} \times \mathbf{l} - mk\frac{\mathbf{r}}{r}, \tag{2.74}$$

donde $\mathbf{l} = \mathbf{r} \times m\dot{\mathbf{r}}$ es el vector de impulso angular.

Para comprobar que este vector es una constante de movimiento simplemente calculamos su derivada temporal

$$\dot{\mathbf{A}} = m\ddot{\mathbf{r}} \times \mathbf{l} + m\dot{\mathbf{r}} \times \dot{\mathbf{l}} - mk\frac{\dot{\mathbf{r}}r - \dot{r}\mathbf{r}}{r^2}, \tag{2.75}$$

pero $\dot{\mathbf{l}} = 0$, $\dot{r} = \mathbf{r} \cdot \dot{\mathbf{r}}/r$ y $m\ddot{\mathbf{r}}$ es la fuerza que vale $-k\mathbf{r}/r^3$, por tanto

$$\dot{\mathbf{A}} = -\frac{km}{r^2}\left[\frac{\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}})}{r} + \dot{\mathbf{r}}r - \frac{\mathbf{r} \cdot \dot{\mathbf{r}}}{r}\mathbf{r}\right], \tag{2.76}$$

y, como $\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}}) = \mathbf{r}\,(\mathbf{r} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}} \cdot (\mathbf{r} \cdot \mathbf{r})$, encontramos que $\dot{\mathbf{A}} = 0$.

El vector de Runge Lenz nos permite encontrar la ecuación de la órbita, veamos esto. El vector de Runge Lenz es perpendicular al vector de impulso angular, por tanto está contenido en el plano de la órbita; esto nos permite tomar su dirección, que desde luego es invariante, como referencia para un sistema de coordenadas polares (r, θ); en este sistema tendremos que

$$\mathbf{A} \cdot \mathbf{r} = Ar\cos\theta, \tag{2.77}$$

pero

$$\mathbf{A} \cdot \mathbf{r} = m\,(\dot{\mathbf{r}} \times \mathbf{l}) \cdot \mathbf{r} - mkr = m\,(\mathbf{r} \times \dot{\mathbf{r}}) \cdot \mathbf{l} - mkr = l^2 - mkr, \tag{2.78}$$

por tanto

$$l^2 - mkr = Ar\cos\theta \tag{2.79}$$

y de aquí encontramos que

$$\frac{1}{r} = \frac{mk}{l^2}\left(1 + \frac{A}{mk}\cos\theta\right) \tag{2.80}$$

que no es otra ecuación que (2.60) cuando la excentricidad es $\varepsilon = A/mk$ y la fase φ es cero. La dirección del vector **A** puede determinarse por el hecho de que la fase en la ecuación de la órbita es cero, pues esto nos hace ver que debe coincidir con el eje de la cónica que pasa por el o los focos. Esta es, quizá, la forma más sencilla de deducir que las órbitas del sistema kepleriano son cónicas.

2.2.3. El péndulo esférico.

Consideremos ahora el péndulo esférico; consiste de una varilla rígida de longitud h y sin masa, uno de cuyos extremos puede girar libremente en torno a un punto fijo del espacio físico tridimensional, encontrándose en el otro extremo una masa m sobre la que actúa una fuerza gravitacional mg.

La partícula se mueve en el espacio tridimensional, pero la varilla constriñe sus posiciones a la superficie bidimensional de una esfera de radio h. Las fuerzas que actúan sobre la masa pueden descomponerse en una componente tangencial a la esfera y una radial; naturalmente la componente radial se combina con el efecto de la rigidez de la varilla para mantener a la partícula sobre la esfera, por esto podemos ver este caso como un caso de dos dimensiones o, mejor dicho, de dos grados de libertad.

Las simetrías del sistema nos hacen ver la conveniencia de utilizar un sistema de coordenadas esféricas cuyo origen coincida con el punto fijo en el espacio físico. Naturalmente la constricción se expresa en la constancia de la coordenada radial, por lo que en realidad tendremos únicamente dos coordenadas (θ, φ).

Aquí debemos tener mucho cuidado, en este sistema el espacio geométrico o físico en que ocurre el movimiento es la superficie de una esfera y en ésta no podemos establecer, con todo propiedad, un sistema de coordenadas. Un sistema de coordenadas no es otra cosa que una correspondencia biunívoca y bicontinua entre los puntos de un espacio y las eneadas de números reales o puntos del espacio R^n, en este sentido sólo se pueden establecer coordenadas cuando el espacio es homeomorfo a R^n y la esfera no lo es. De manera similar debemos decir que coordenadas como las polares para el plano o las esféricas para el espacio tridimensional no son, estrictamente hablando, coordenadas, ya que la correspondencia que establecen no es biunívoca. Esto introduce dificultades que únicamente se pueden salvar desarrollando la teoría de las

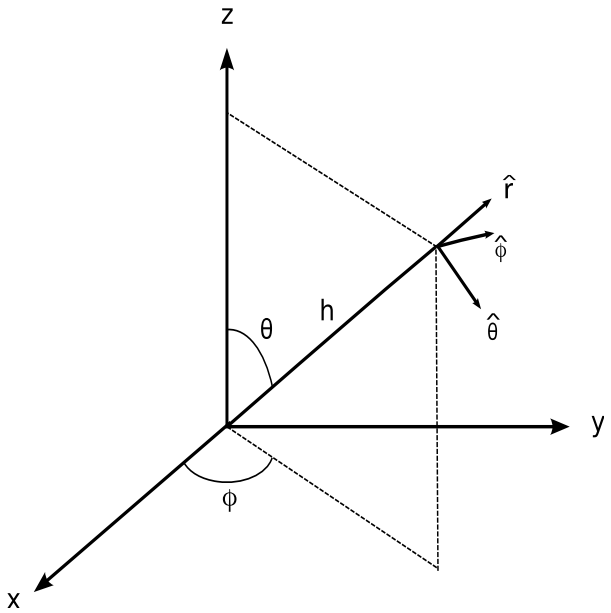


Figura 2.12: Vectores de la base unitaria inducida por las coordenadas esféricas.

variedades diferenciales y el cálculo sobre dichas variedades. Sin embargo en casos simples, como el que estamos estudiando, podemos salvar dichas dificultades utilizando sistemas de "coordenadas", como las esféricas, siempre y cuando seamos plenamente conscientes de lo que se hace y se utilicen argumentos sólidos para darle la vuelta a las dificultades e inconsistencias que puedan aparecer, como hicimos en el caso del péndulo simple.

Ahora en coordenadas esféricas ocurre algo similar a lo que ocurre en polares y en general en los sistemas curvilíneos de coordenadas, los vectores que conforman una base inducida por las coordenadas cambian de punto en punto y por tanto al expresar las ecuaciones vectoriales de movimiento en coordenadas, debemos tomar en consideración la variación de dicha base.

En coordenadas esféricas (r, θ, φ) la base unitaria inducida por las coordenadas son los vectores $(\hat{\mathbf{r}}, \hat{\theta}, \hat{\varphi})$ definidos en la figura 2.12.

Podemos ver de esta figura que las parciales de estos vectores con respecto

a la coordenada r son cero y que

$$\frac{\partial \hat{\mathbf{r}}}{\partial \theta} = \hat{\theta}, \tag{2.81}$$

$$\frac{\partial \hat{\theta}}{\partial \theta} = -\hat{\mathbf{r}}, \tag{2.82}$$

$$\frac{\partial \hat{\varphi}}{\partial \theta} = 0, \tag{2.83}$$

$$\frac{\partial \hat{\mathbf{r}}}{\partial \varphi} = \hat{\varphi}\text{sen}\theta, \tag{2.84}$$

$$\frac{\partial \hat{\theta}}{\partial \varphi} = \hat{\varphi}\cos\theta, \tag{2.85}$$

$$\frac{\partial \hat{\varphi}}{\partial \varphi} = -\hat{\mathbf{r}}\text{sen}\theta - \hat{\theta}\cos\theta. \tag{2.86}$$

Por tanto

$$\frac{d\mathbf{r}}{dt} = \dot{r}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\theta} + r\dot{\varphi}\text{sen}\theta\hat{\varphi}, \tag{2.87}$$

y tomando en cuenta que en nuestro caso $r = h$ es constante y únicamente se debe tomar en consideración la velocidad tangencial

$$\frac{d\mathbf{r}}{dt} = h\dot{\theta}\hat{\theta} + h\dot{\varphi}\text{sen}\theta\hat{\varphi}. \tag{2.88}$$

Tomando segunda derivada y nuevamente considerando que $\dot{r} = 0$ y que únicamente estamos interesados en la aceleración tangencial obtenemos que

$$\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{tan} = h(\ddot{\theta} - \dot{\varphi}^2\text{sen}\theta\cos\theta)\hat{\theta} + h(\ddot{\varphi}\text{sen}\theta + 2\dot{\theta}\dot{\varphi}\cos\theta)\hat{\varphi}. \tag{2.89}$$

La fuerza de gravedad es $mg(-\cos\theta\hat{\mathbf{r}} + \text{sen}\theta\hat{\theta})$, pero debemos considerar únicamente la componente tangencial $mg\text{sen}\theta\hat{\theta}$. Por tanto la ecuación de movimiento en coordenadas (θ, φ) se expresará como

$$mh\left(\ddot{\theta} - \dot{\varphi}^2\text{sen}\theta\cos\theta\right) = m \quad g\text{sen}\theta, \tag{2.90}$$

$$mh\left(\ddot{\varphi}\text{sen}\theta + 2\dot{\theta}\dot{\varphi}\cos\theta\right) = 0. \tag{2.91}$$

Multiplicando la segunda de estas ecuaciones por $\text{sen}\theta$ vemos que se reduce a

$$mh\frac{d}{dt}(\dot{\varphi}\text{sen}^2\theta) = 0 \tag{2.92}$$

y por tanto existe una integral de movimiento adicional a la energía total. Esta integral de movimiento la podemos identificar con la proyección del impulso angular en la dirección vertical

$$p_\varphi = mh^2\dot{\varphi}\text{sen}^2\theta. \tag{2.93}$$

Despejando $\dot{\varphi}$ y substituyendo en la primera de las ecuaciones (2.91) obtenemos

$$mh\ddot{\theta} = \frac{p_\varphi^2 \cos\theta}{mh^3\text{sen}^3\theta} + mg\text{sen}\theta. \tag{2.94}$$

Esta ecuación puede ser reducida a cuadraturas como toda ecuación en una sola variable y obtener $\theta(t)$, aunque la integral resultante no puede expresarse en términos de funciones elementales.

Una vez que conocemos $\theta(t)$ hacemos uso de la ec. (2.93) para obtener $\varphi(t)$. En estos dos procesos se generan dos constantes de integración que, aunadas a p_φ y E, constituyen las cuatro condiciones iniciales requeridas y el problema del movimiento del péndulo esférico queda reducido a cuadraturas.

Ahora bien al observar la ec. (2.94) nos damos cuenta que es equivalente a la del movimiento de una partícula sujeta a un potencial efectivo de la forma

$$V_{ef} = \frac{p_\varphi^2}{2mh^2\text{sen}^2\theta} + mgh\cos\theta, \tag{2.95}$$

esto nos permite utilizar nuevamente las técnicas del análisis cualitativo que hemos estado usando. Hay que señalar que la introducción de un factor h se hace necesaria debido a que en coordenadas esféricas el gradiente de una función $f(r,\theta,\varphi)$ queda expresado por

$$\text{grad}f(r,\theta,\varphi) = \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial\theta}\hat{\theta} + \frac{1}{r\text{sen}\theta}\frac{\partial f}{\partial\varphi}\hat{\varphi}. \tag{2.96}$$

En la figura 2.13 se presenta la gráfica de este potencial efectivo.

Podemos ver que la variación en θ estará siempre acotada ya que el potencial efectivo se va a infinito para $\theta = 0$ y para $\theta = \pi$. Como la función tiene un solo mínimo, la variable θ siempre oscilará periódicamente. La ec. (2.93) nos proporciona el comportamiento de la variable φ y podemos ver que, si bien la velocidad angular en φ varía con el cambio en θ, el péndulo girará monótonamente en torno a la dirección vertical en un sentido u otro dependiendo del signo de p_φ, o sea de las condiciones iniciales. Así el movimiento general del péndulo esférico será el de una variación periódica del ángulo que hace con la vertical, combinada con un giro en torno de esa misma dirección. Ocurre aquí algo similar a lo ya señalado para el movimiento

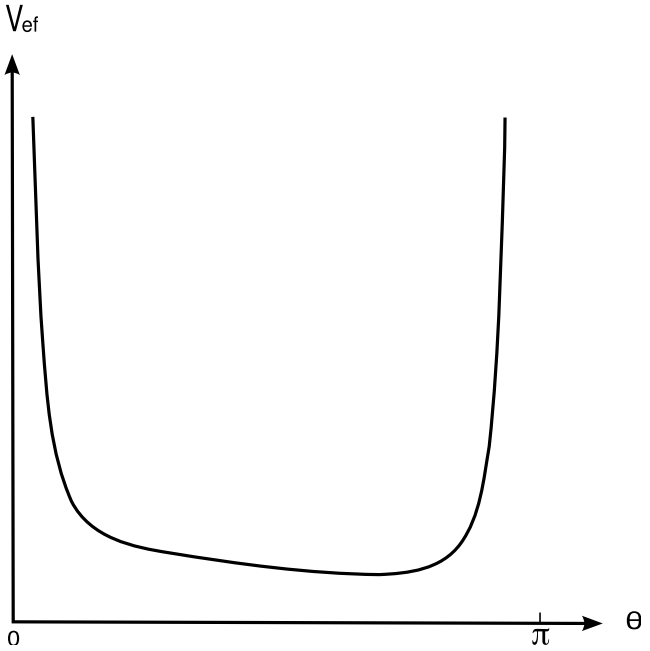


Figura 2.13: Potencial efectivo del péndulo esférico.

en un campo central, puede ocurrir que el período del movimiento vertical sea conmensurable con el período de las rotaciones, esto es que el péndulo suba y baje un cierto número entero de veces en el tiempo en que realiza un número entero de vueltas, en este caso el movimiento total será periódico; puede también ocurrir que está coincidencia no se dé, en ese caso el movimiento será cuasiperiódico y la posición de la partícula cubrirá densamente una banda en la esfera comprendida entre los círculos correspondientes a los valores máximo y mínimo de θ; el que tengamos un caso o el otro dependerá de las condiciones iniciales. La figura 2.14 muestra el aspecto general del movimiento del péndulo esférico.

Puede verse que la posición del mínimo está siempre comprendida en el intervalo $(\pi/2, \pi)$ y se acerca al primero cuando p_φ tiende a infinito. Naturalmente si la energía total coincide con el valor del potencial efectivo en el mínimo, el ángulo θ será constante y la trayectoria del péndulo trazará un círculo horizontal. Es interesante señalar que cuando $p_\varphi = 0$ el péndulo esférico se reduce al simple y cuando $p_\varphi \to \infty$ el péndulo tiende a trazar un círculo máximo horizontal. Es también interesante hacer ver que si $p_\varphi \neq 0$, esto es, si tenemos un péndulo esférico que no se reduce a uno simple, la partícula

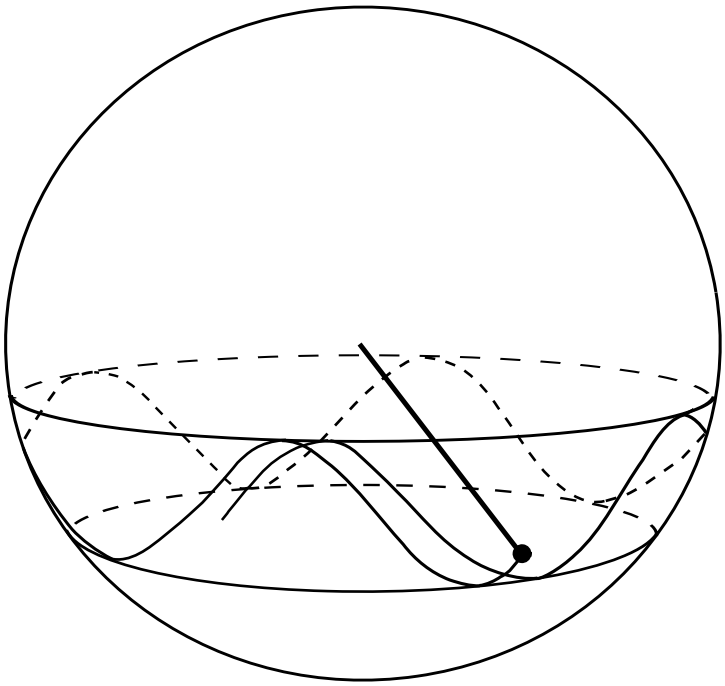


Figura 2.14: El movimiento del péndulo esférico.

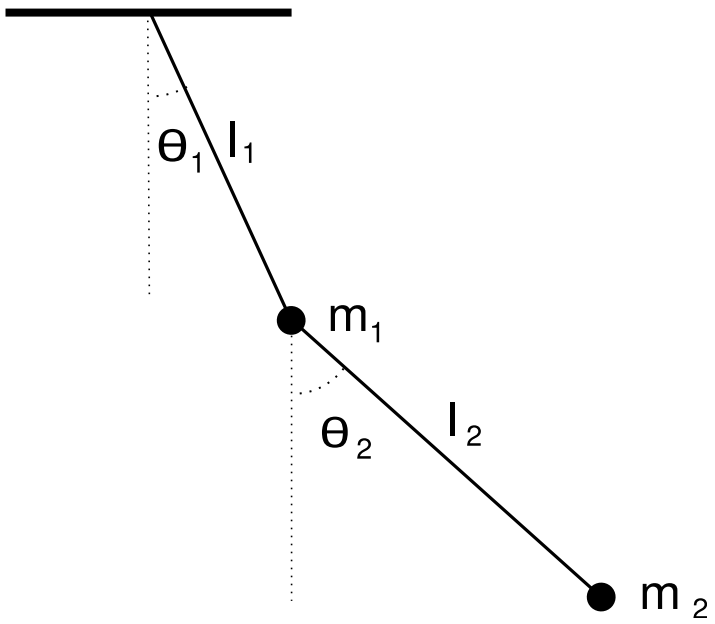


Figura 2.15: El péndulo doble.

no puede pasar por los puntos $\theta = 0$ o $\theta = \pi$, que resultan ser aquellos en que las coordenadas esféricas presentan singularidades, como el movimiento no pasa por ellos, no constituyen entonces una mayor preocupación.

2.2.4. El péndulo doble.

Pasemos a estudiar el llamado péndulo doble. Consiste en un péndulo simple de cuya masa cuelga otro péndulo, ver la figura 2.15.

En este caso no tenemos una sola partícula moviéndose en un espacio o variedad de dos dimensiones, tenemos dos partículas, una primera moviéndose en un círculo, esto es una variedad de una dimensión, y una segunda que, para cada posición de la primera, se mueve sobre un círculo centrado en esa primera partícula.

Podemos ver de la figura que la situación geométrica total del sistema queda determinada por dos ángulos, θ_1 y θ_2, los que podemos tomar como coordenadas de un espacio bidimensional. Este espacio se genera trazando un círculo por cada punto de otro círculo, se trata pues del producto cartesiano de dos círculos o esferas de una dimensión $S^1 \times S^1$. Análogamente un plano se puede ver como el producto cartesiano de dos rectas o, en términos pictóricos, el espacio que se genera al deslizar una recta sobre otra; podemos

ver que al deslizar un círculo sobre otro círculo se genera un toro de dos dimensiones, esto es, la variedad $S^1 \times S^1$ es un toro de dos dimensiones y podemos considerar a los ángulos θ_1 y θ_2 como "coordenadas" para el toro; desde luego hemos puesto coordenadas entre comillas pues en sentido estricto no existe un sistema de coordenadas para el toro y se aplican a este caso los mismos comentarios que hicimos para el caso de la esfera. Así el péndulo doble puede estudiarse como el movimiento de un punto sobre la superficie de un toro bidimensional y, en este sentido, es un sistema mecánico en dos dimensiones o mejor dicho en dos grados de libertad.

Para obtener las ecuaciones de movimiento debemos hacerlo partícula por partícula tomando en consideración las fuerzas que se ejercen sobre cada una de ellas, incluyendo, desde luego, las que ejercen las varillas. La liga entre ellas establecida por la varilla que las une y la influencia de la varilla que sujeta a la primera, establece una especie de interacción entre ellas, esto es, la fuerza que hay que considerar sobre cada una de ellas puede verse como la de la gravedad más una de interacción. Finalmente obtendremos las ecuaciones diferenciales de movimiento en θ_1 y θ_2.

Ahora bien este caso es más complejo que el anterior y hacer los cálculos que nos permitan establecer las ecuaciones de movimiento es más laborioso. Pero adicionalmente a esta dificultad aparece otra de mayor trascendencia, ya que, si bien la energía total es una integral de movimiento por ser la fuerza de gravedad conservativa, no podemos encontrar otra integral de movimiento independiente de ella. Esto nos impide reducir el problema a una dimensión y utilizar el método de análisis cualitativo que hemos venido usando; es más, nos impide reducir a cuadraturas las ecuaciones. Esta cuestión es de mucha más profundidad que la simple complejidad del procedimiento para encontrar las ecuaciones de movimiento; es una dificultad esencial, no se trata de que sea más o menos difícil encontrar la segunda integral de movimiento, se puede demostrar que no existe y que eso impide definitivamente la reducción a cuadraturas.

Casos como éste se denominan no integrables y en general presentan comportamientos dinámicos sumamente complejos y sus trayectorias o trayectorias resultan ser, en muchas ocasiones, curvas que dan la apariencia de ser totalmente irregulares o caóticas. Estas trayectorias forman parte del fenómeno conocido como caos dinámico descubierto a fines de los años sesenta del siglo XX y que constituye una de las áreas de investigación más importantes dentro de la teoría de los sistemas dinámicos y de los sistemas mecánicos en particular. El estudio de estos fenómenos mecánicos requiere del desarrollo de conceptos, métodos y procedimientos que constituyen la parte medular de la mecánica analítica, por lo que será tratado en los últimos

capítulos de este libro.

2.2.5. El cuerpo rígido libre.

Estudiemos ahora un sistema de tres dimensiones, el cuerpo rígido libre, esto es aquél sobre el que no actúan fuerzas.

Para establecer totalmente la posición de un cuerpo rígido podemos determinar primero la posición de uno de sus puntos y luego, por medio de una rotación con respecto a ese punto, la situación del cuerpo. Para determinar la posición del primer punto requerimos de tres coordenadas y para determinar la rotación bastan tres parámetros adicionales. Para un cuerpo totalmente libre podemos utilizar como punto de referencia al centro de masas del cuerpo y considerarlo como un punto fijo en el espacio. Este es un caso especial del estudio del movimiento de un cuerpo rígido cuando uno de cuyos puntos se mantiene fijo, por ello plantearemos este sistema más general.

Pongamos un sistema de referencia en el punto fijo y un sistema de coordenadas cartesianas (x, y, z) cuyo origen coincida con dicho punto fijo, llamaremos a este sistema el sistema inercial. Nuestro estudio se reduce entonces al de las rotaciones del cuerpo con respecto a este sistema, para lo cual requerimos únicamente de tres parámetros.

El cuerpo rígido lo entendemos como una distribución rígida y acotada de masa que queda definida por una densidad de masa ρ. La masa total será

$$M = \int\int\int \rho dV \tag{2.97}$$

donde la integral se hace sobre el volumen que ocupa el cuerpo rígido. En el sistema de coordenadas que hemos escogido tendremos

$$M = \int\int\int \rho(x, y, z) dx dy dz. \tag{2.98}$$

La energía cinética del elemento de volumen en la posición $\mathbf{r}$ será

$$\frac{1}{2} (\mathbf{v}_r \cdot \mathbf{v}_r) \, dm = \frac{1}{2} (\mathbf{v}_r \cdot \mathbf{v}_r) \, \rho dV \tag{2.99}$$

donde $\mathbf{v}_r$ es el vector velocidad del punto que se encuentra en la posición $\mathbf{r}$. Como el movimiento del cuerpo es una rotación en torno del punto fijo

$$\mathbf{v}_r = \boldsymbol{\omega} \times \mathbf{r}, \tag{2.100}$$

donde ω en el vector instantáneo de velocidad angular; tendremos que la energía cinética total es

$$T = \frac{1}{2}\int\int\int (\boldsymbol{\omega}\times\mathbf{r})\cdot(\boldsymbol{\omega}\times\mathbf{r})\rho dV \tag{2.101}$$

$$= \frac{1}{2}\int\int\int [r^2(\boldsymbol{\omega}\cdot\boldsymbol{\omega}) - (\mathbf{r}\cdot\boldsymbol{\omega})^2]\rho dV, \tag{2.102}$$

y en coordenadas cartesianas

$$T = \frac{1}{2}\int\int\int [(x^2+y^2+z^2)(\omega_x^2+\omega_y^2+\omega_z^2) - (x\omega_x + y\omega_y + z\omega_x)^2]\rho dxdydz. \tag{2.103}$$

Si definimos, utilizando notación diádica, el tensor simétrico $\mathcal{I}$ por

$$\mathcal{I} = \int\int\int [(\mathbf{r}\cdot\mathbf{r})\mathbf{1} - \mathbf{rr}]\rho dV \tag{2.104}$$

o en coordenadas cartesianas

$$\mathcal{I}_{xx} = \int\int\int (y^2+z^2)\rho dxdydz, \tag{2.105}$$

$$\mathcal{I}_{yy} = \int\int\int (x^2+z^2)\rho dxdydz, \tag{2.106}$$

$$\mathcal{I}_{zz} = \int\int\int (x^2+y^2)\rho dxdydz, \tag{2.107}$$

$$\mathcal{I}_{xy} = \mathcal{I}_{yx} = \int\int\int xy\rho dxdydz, \tag{2.108}$$

$$\mathcal{I}_{xz} = \mathcal{I}_{zx} = \int\int\int xz\rho dxdydz, \tag{2.109}$$

$$\mathcal{I}_{yz} = \mathcal{I}_{zy} = \int\int\int yz\rho dxdydz, \tag{2.110}$$

tendremos que

$$T = \frac{1}{2}\mathcal{I}[\boldsymbol{\omega},\boldsymbol{\omega}] = \frac{1}{2}\sum_{ij}\mathcal{I}_{ij}\omega_i\omega_j, \tag{2.111}$$

donde los índices i y j toman los valores x, y y z.

Este tensor es el tensor de inercia del cuerpo rígido con respecto al punto fijo, que, como podemos ver, es un tensor simétrico y positivo definido.

El impulso angular $\mathbf{l}_r$ del elemento de masa en el punto $\mathbf{r}$ será

$$\mathbf{l}_r = \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})\rho dV. \tag{2.112}$$

El impulso angular total con respecto al punto fijo será entonces

$$\mathbf{l} = \int\int\int \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})\rho dV = \int\int\int [r^2\boldsymbol{\omega} - \mathbf{r}(\mathbf{r} \cdot \boldsymbol{\omega})]dV, \tag{2.113}$$

que queda expresado en términos del tensor de inercia por

$$\mathbf{l} = \mathcal{I}[\bullet, \boldsymbol{\omega}] = \left(\sum_j \mathcal{I}_{xj}\omega_j, \sum_j \mathcal{I}_{yj}\omega_j, \sum_j \mathcal{I}_{zj}\omega_j \right). \tag{2.114}$$

Debemos señalar que al estar el tensor de inercia referido al sistema inercial será dependiente del tiempo, ya que el cuerpo rígido está en movimiento respecto a ese sistema.

En este sistema inercial las ecuaciones de movimiento son

$$\frac{d}{dt}\mathbf{l} = \frac{d}{dt}\left(\mathcal{I}[\bullet, \boldsymbol{\omega}]\right) = \mathbf{n}, \tag{2.115}$$

donde $\mathbf{n}$ es la torca de las fuerzas que actúan sobre el cuerpo, con respecto al punto fijo.

A partir de las ecuaciones (2.115) es muy laborioso obtener unas ecuaciones en componentes. Esto se debe a que el tensor de inercia en el sistema inercial no es constante, como si lo es la masa de una partícula en la segunda ley de Newton. En adición a esto la forma de esta dependencia temporal está en función del movimiento del cuerpo, mismo que queremos obtener a partir de las ecuaciones. Otro tanto podemos decir respecto a la torca. Esta dificultad se relaciona con el hecho de que no hemos establecido unas coordenadas que determinen la posición del cuerpo rígido. El sistema inercial (x, y, z) determina la posición de los puntos que conforman el cuerpo rígido, pero no son tres parámetros que determinan totalmente la posición del cuerpo, en términos de los cuales pudiéramos expresar el tensor de inercia y los vectores de impulso angular, velocidad angular, torca y finalmente las ecuaciones (2.115).

No seguiremos aquí la construcción de un sistema de coordenadas para el cuerpo rígido, cuestión que dejamos para más adelante, atacaremos el problema del cuerpo rígido libre por otra vía.

Decíamos que una dificultad reside en la dependencia temporal del tensor de inercia calculado respecto al sistema inercial. Si calculamos este mismo

tensor en un sistema de referencia anclado al cuerpo rígido esta dificultad desaparece puesto que en él el tensor sería constante; en vez de una dependencia temporal del tensor de inercia tendríamos un sistema no inercial. Tomemos, pues, un sistema de referencia con origen en el mismo punto fijo pero anclado al cuerpo rígido y un sistema cartesiano de coordenadas en él (X, Y, Z). Así como hemos utilizado letras mayúsculas para expresar las coordenadas en este sistema, utilizaremos letras mayúsculas para expresar los vectores en él, esto es el impuso angular total en el sistema anclado al cuerpo será $\mathbf{L}$, la torca $\mathbf{N}$, la velocidad angular $\mathbf{\Omega}$ y el tensor de inercia $\mathbf{I}$.

En cada instante del tiempo t la relación entre el sistema inercial y el anclado al cuerpo será una rotación en tres dimensiones dada por un operador lineal, ortogonal y que preserva orientación R_t, tendremos que

$$\mathbf{l} = R_t[\mathbf{L}], \;\; \boldsymbol{\omega} = R_t[\mathbf{\Omega}], \;\; \mathbf{n} = R_t[\mathbf{N}]. \tag{2.116}$$

El tensor de inercia con respecto a un sistema anclado al cuerpo es constante, simétrico y positivo definido, por tanto existe un sitema de coordenadas en que queda diagonal con todos los términos diagonales positivos. Los ejes coordenados de un sistema así están en la dirección de los eigenvectores del tensor y los valores en la diagonal serían los eigenvalores correspondientes[5]. Supondremos siempre que el sistema (X, Y, Z) anclado al cuerpo es un sistema de esta naturaleza. En él el tensor de inercia es constante y queda expresado de manera diagonal con

$$I_{XX} = \int\int\int (Y^2 + Z^2)\rho dX dY dZ, \tag{2.117}$$

$$I_{YY} = \int\int\int (X^2 + Z^2)\rho dX dY dZ, \tag{2.118}$$

$$I_{ZZ} = \int\int\int (X^2 + Y^2)\rho dX dY dZ. \tag{2.119}$$

Para aprovechar la constancia del tensor de inercia en el sistema anclado al cuerpo, traslademos la ecuación vectorial 2.115 a ese sistema de referencia. Calculemos primero $d\mathbf{l}/dt$ en el sistema del cuerpo. Tenemos que

$$\mathbf{l} = R_t[\mathbf{L}], \tag{2.120}$$

entonces

$$\frac{d}{dt}\mathbf{l} = \dot{R}_t[\mathbf{L}] + R_t[\dot{\mathbf{L}}] = \dot{R}_t[R_t^{-1}[\mathbf{l}]] + R_t[\dot{\mathbf{L}}] = \dot{R}_t R_t^{-1}[\mathbf{l}] + R_t[\dot{\mathbf{L}}]. \tag{2.121}$$

[5]Estas tres direcciones determinan los llamados ejes principales de inercia del cuerpo rígido con respecto al punto fijo.

Ahora R_t es un operador ortogonal por lo que su inversa es su transpuesta, esto es

$$R_t R_t^T = e, \tag{2.122}$$

por lo que

$$\dot{R}_t R_t^T + R_t \dot{R}_t^T = 0, \tag{2.123}$$

esto implica que

$$\dot{R}_t R_t^T = -R_t \dot{R}_t^T = -\left(\dot{R}_t R_t^T\right)^T, \tag{2.124}$$

esto es $\dot{R}_t R_t^T$ o lo que es lo mismo $\dot{R}_t R_t^{-1}$ es un operador antisimétrico. Sabemos que en tres dimensiones aplicar un operador antisimétrico a un vector cualquiera es lo mismo que multiplicarlo vectorialmente por un vector asociado al operador, que en nuestro caso, por ser R_t una rotación se trata del vector de velocidad angular. Así tendremos que

$$\dot{R}_t R_t^{-1}[\mathbf{l}] = \boldsymbol{\omega} \times \mathbf{l} \tag{2.125}$$

y

$$\frac{d}{dt}\mathbf{l} = \boldsymbol{\omega} \times \mathbf{l} + R_t[\dot{\mathbf{L}}]. \tag{2.126}$$

Ahora las rotaciones preservan el producto vectorial por lo que

$$\frac{d}{dt}\mathbf{l} = R_t[\boldsymbol{\Omega}] \times R_t[\mathbf{L}] + R_t[\dot{\mathbf{L}}] = R_t\left[\boldsymbol{\Omega} \times \mathbf{L} + \dot{\mathbf{L}}\right]. \tag{2.127}$$

Con respecto a la torca tenemos que $\mathbf{n} = R_t[\mathbf{N}]$, por lo que la ecuación (2.115) queda como

$$R_t[\mathbf{N}] = R_t\left[\boldsymbol{\Omega} \times \mathbf{L} + \dot{\mathbf{L}}\right], \tag{2.128}$$

lo que implica

$$\mathbf{N} = \boldsymbol{\Omega} \times \mathbf{L} + \dot{\mathbf{L}} \tag{2.129}$$

y utilizando el tensor de inercia tendremos que

$$\mathbf{N} = \boldsymbol{\Omega} \times \mathbf{I}[\bullet, \Omega] + \frac{d}{dt}\left(\mathbf{I}[\bullet, \Omega]\right), \tag{2.130}$$

que por ser el tensor de inercia en el cuerpo constante se reduce a

$$\mathbf{N} = \boldsymbol{\Omega} \times \mathbf{I}[\bullet, \boldsymbol{\Omega}] + \mathbf{I}[\bullet, \dot{\boldsymbol{\Omega}}] \tag{2.131}$$

que es la ecuación de movimiento expresada en el sistema del cuerpo. En las coordenadas cartesianas (X, Y, Z) en las que el tensor es diagonal tendremos

$$N_X = \Omega_Y \Omega_Z (I_{ZZ} - I_{YY}) + I_{XX} \dot{\Omega}_X, \tag{2.132}$$

$$N_Y = \Omega_X \Omega_Z (I_{XX} - I_{ZZ}) + I_{YY} \dot{\Omega}_Y, \tag{2.133}$$

$$N_Z = \Omega_X \Omega_Y (I_{YY} - I_{XX}) + I_{ZZ} \dot{\Omega}_Z. \tag{2.134}$$

Estas ecuaciones se conocen como las ecuaciones de Euler para el cuerpo rígido, sus soluciones nos proporcionan el vector de velocidad angular expresado en el sistema anclado al cuerpo.

Naturalmente es necesario tener la torca expresada en el sistema anclado al cuerpo, lo que muchas veces no resulta sencillo y en general obtendremos unas ecuaciones de muy difícil solución.

En el caso del cuerpo rígido libre la torca es cero, tanto en el sistema inercial como en el anclado al cuerpo, por lo que en ese caso tendremos

$$I_{XX} \dot{\Omega}_{XX} + \Omega_Y \Omega_Z (I_{ZZ} - I_{YY}) = 0, \tag{2.135}$$

$$I_{YY} \dot{\Omega}_{YY} + \Omega_X \Omega_Z (I_{XX} - I_{ZZ}) = 0, \tag{2.136}$$

$$I_{ZZ} \dot{\Omega}_{ZZ} + \Omega_X \Omega_Y (I_{YY} - I_{XX}) = 0, \tag{2.137}$$

que, por ser las I constantes, constituyen una ecuación diferencial completa autónoma que podemos pasar a resolver y encontrar así la evolución temporal de las componentes de la velocidad angular en el sistema anclado al cuerpo.

Las ecuaciones (2.136) implican la existencia de dos integrales de movimiento

$$E = \frac{1}{2} \left(I_X \Omega_X^2 + I_Y \Omega_Y^2 + I_Z \Omega_Z^2 \right) \tag{2.138}$$

y

$$L^2 = I_X^2 \Omega_X^2 + I_Y^2 \Omega_Y^2 + I_Z^2 \Omega_Z^2. \tag{2.139}$$

La primera es la energía total y la segunda el cuadrado del impulso angular total, éstas se obtienen de su expresión en el sistema inercial y de la invariancia del producto escalar ante rotaciones o bien de la aplicación directa de las ecuaciones de Euler. A partir de estas integrales de movimiento podemos poner, digamos, Ω_Y y Ω_Z en función de Ω_X, substituir en las ecuaciones de Euler, obtener una ecuación de primer orden para Ω_X y resolverla, pero no seguiremos este camino, en su lugar seguiremos un proceso geométrico para visualizar el movimiento.

Las integrales de movimiento se pueden expresar como

$$E = \frac{1}{2} \left(\frac{L_X^2}{I_X} + \frac{L_Y^2}{I_Y} + \frac{L_Z^2}{I_Z} \right) \tag{2.140}$$

y

$$L^2 = L_X^2 + L_Y^2 + L_Z^2, \tag{2.141}$$

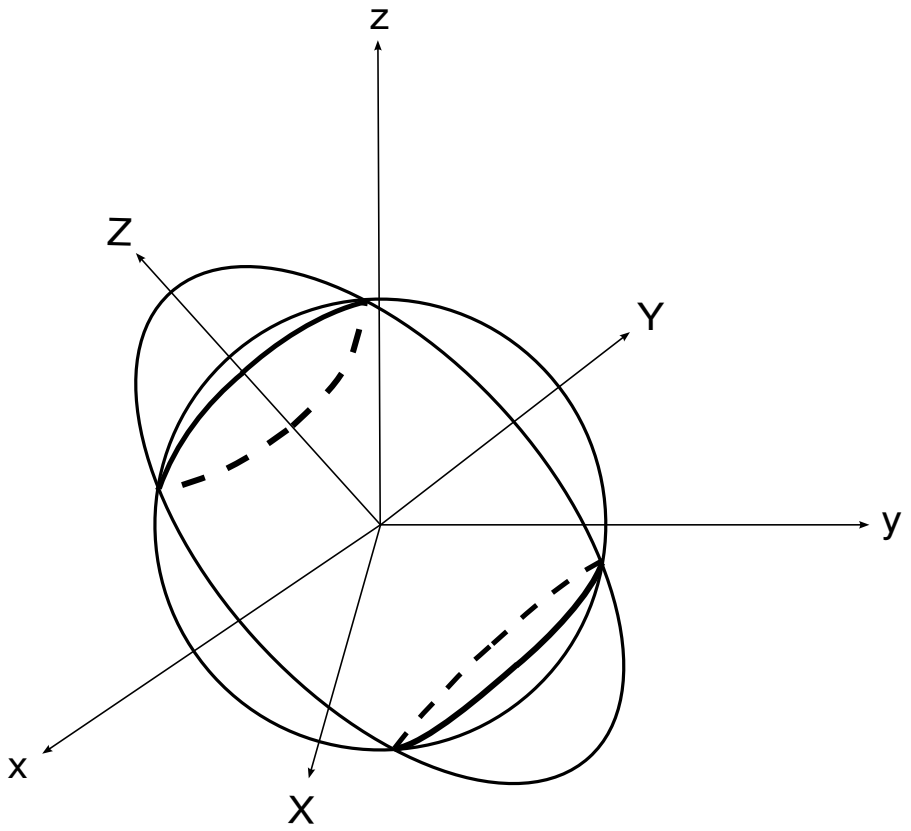


Figura 2.16: El vector de impulso angular en el sistema del cuerpo taza las curvas que resultan de la intersección de una esfera con un elipsoide.

por lo que el extremo del vector de impulso angular con respecto al sistema del cuerpo está sobre la superficie del elipsoide dado por la primera ecuación y sobre una esfera especificada por la segunda, esto es, debe permanecer sobre la curva en que estas dos superficies se intersectan como se muestra en la figura 2.16[6].

Los semiejes del elipsoide valen $\sqrt{2EI_X}$, $\sqrt{2EI_Y}$ y $\sqrt{2EI_Z}$, por lo que la magnitud del impulso angular no podrá ser mayor que el mayor de ellos ni menor que el más pequeño. Ahora, cuando la magnitud del impulso angular coincida con alguno de estos valores, la intersección se reduce a dos puntos diametralmente opuestos que son los puntos sobre el elipsoide donde

[6]Debemos recordar que el vector de impulso angular desde la perspectiva del sistema anclado al cuerpo no permanece constante.

lo intersecta el eje correspondiente, por ello el vector **L** será constante; esto representa, con respecto al sistema inercial, una rotación del cuerpo rígido en torno al eje principal de inercia correspondiente, este movimiento es estacionario ya que el eje de rotación esta fijo en el espacio y la velocidad angular es constante. Cuando la magnitud del impulso angular coincida con el valor del semieje intermedio, la intersección estará compuesta de dos curvas que se intersectan en los puntos donde el eje intermedio corta al elipsoide, si la magnitud del impuso angular coincide con el valor del semieje intermedio, tendremos también una rotación estacionaria, en este caso en torno al eje intermedio. Ahora ante una pequeña perturbación de estos movimientos estacionarios obtenemos dos comportamientos distintos. Es fácil ver que si se trata del semieje mayor o del menor del elipsoide, el impulso angular en el cuerpo trazará una pequeña curva cerrada en torno al punto estacionario, o mejor dicho, el impulso angular realizará una pequeña precesión en torno al semieje correspondiente del elipsoide; con respecto al sistema inercial, el cuerpo tendrá un pequeño bamboleo en adición a una rotación en trono al eje principal de inercia correspondiente. En caso de que se trate del valor intermedio, ante una pequeña perturbación el impulso angular trazará una curva sobre el elipsoide que lo aleja considerablemente del punto estacionario; esto es, con respecto al sistema inercial, el cuerpo, en adición a la rotación en torno a al eje principal de inercia, presentará fuertes movimientos de bamboleo. Vemos entonces que de los tres movimientos estacionarios de rotación de un cuerpo rígido libre, dos, los correspondientes a las rotaciones en torno a los ejes principales con momentos de inercia mayor y menor son estables y el tercero, el correpondiente a una rotación en torno al tercer eje principal, aquel con momento de inercia intermedio, es inestable. El lector puede sentir esto fácilmente tratando de lanzar una raqueta al aire al tiempo que le imprime una rotación, ya sea en torno al eje definido por el mango, al eje perpendicular a la pala o al perpendicular al mago contenido en la pala; en los primeros caso es fácil hacerlo sin producir fuertes bamboleos, en el tercero es prácticamente imposible.

El movimiento general del cuerpo respecto al sistema inercial se puede visualizar por medio del procedimiento llamado de Poinsot. Dado un cuerpo rígido y un punto en él, la superficie trazada por un vector de la forma

$$\frac{\mathbf{e}}{\sqrt{I_e}}, \tag{2.142}$$

donde **e** es un vector unitario que traza la esfera unitaria centrada en dicho punto y I_e es el momento de inercia del cuerpo respecto al eje determinado por **e**, es un elipsoide cuyo centro coincide con el punto de referencia. Esto se

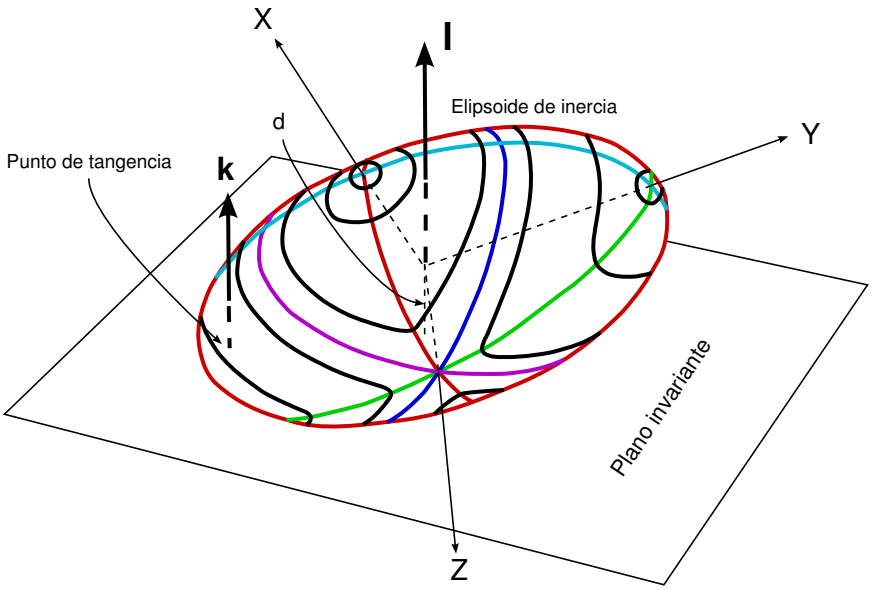


Figura 2.17: En la construcción de Poisont el elipsoide de inercia con respecto al punto fijo rueda sin resvalar sobre un plano invariante perpendicular al impulso angular y a una distancia d del centro del elipsoide. Las curvas que se trazan sobre el elipsoide se llaman las polodas, de las que se han dibujado algunas, las que quedan trazadas sobre el plano se conocen como herpolodas.

debe a que el tensor de inercia es una forma cuadrática positiva definida, por lo que la superficie trazada es cuadrática y positiva, esto es un elipsoide. A este elipsoide se le conoce como el elipsoide de inercia del cuerpo rígido con respecto al punto de referencia. En nuestro caso y con respecto al sistema anclado al punto fijo la ecuación del elipsoide de inercia será

$$I_{XX}X^2 + I_{YY}Y^2 + I_{ZZ}Z^2 = 1, \tag{2.143}$$

ya que los ejes del elipsoide de inercia coinciden con los ejes principales de inercia y con los ejes coordenados.

En un instante del tiempo, el eje trazado sobre la dirección de la velocidad angular $\mathbf{\Omega}$ intersecta al elipsoide de inercia en dos puntos, llamemos $\mathbf{R}$ al vector de posición de uno de esos puntos, ver la figura 2.17. Como este punto está sobre el elipsoide

$$|\mathbf{R}| = \frac{1}{\sqrt{I_R}}, \tag{2.144}$$

donde I_R es el momento de inercia con respecto al eje definido por $\mathbf{R}$. La

energía total, que es únicamente cinética, es

$$T = \frac{1}{2} I[\mathbf{\Omega}, \mathbf{\Omega}], \tag{2.145}$$

que será igual a

$$T = \frac{1}{2} \frac{|\mathbf{\Omega}|^2}{|\mathbf{R}|^2} I[\mathbf{R}, \mathbf{R}], \tag{2.146}$$

pero como $\mathbf{R}$ está sobre el elipsoide $I[\mathbf{R}, \mathbf{R}] = 1$, por tanto

$$T = \frac{1}{2} \frac{|\mathbf{\Omega}|^2}{|\mathbf{R}|^2}. \tag{2.147}$$

Consideremos ahora un plano perpendicular al vector de impulso angular que sea tangente al elipsoide de inercia, en el punto de tangencia la normal al elipsoide será paralela al vector de impulso angular, por serlo al plano, por lo que, si (X, Y, Z) son las coordenadas del punto de tangencia, la normal al elipsoide es

$$(2I_{XX}X, 2I_{YY}Y, 2I_{ZZ}Z) \tag{2.148}$$

que debe ser proporcional al vector $\mathbf{L}$, por tanto

$$2\left(I_{XX}X, I_{YY}Y, I_{ZZ}Z\right) = \alpha\left(I_{XX}\Omega_X, I_{YY}\Omega_Y, I_{ZZ}\Omega_Z\right), \tag{2.149}$$

donde α es la constante de proporcionalidad. Esta relación muestra que

$$X = \frac{\alpha}{2}\Omega_X, \;\; Y = \frac{\alpha}{2}\Omega_Y, \;\; Z = \frac{\alpha}{2}\Omega_Z, \tag{2.150}$$

y utilizando las relación general (2.147) obtenemos que $\alpha/2 = 1/\sqrt{2T}$, esto es el punto de tangencia está sobre el eje instantáneo de rotación y su vector de posición es

$$\mathbf{K} = \left(\frac{\Omega_X}{\sqrt{2T}}, \frac{\Omega_Y}{\sqrt{2T}}, \frac{\Omega_Z}{\sqrt{2T}}\right). \tag{2.151}$$

La distancia d del centro del elipsoide al plano estará dada por

$$d = \frac{\mathbf{k} \cdot \mathbf{L}}{|\mathbf{L}|} = \frac{\left(\frac{\Omega_X}{\sqrt{2T}}, \frac{\Omega_Y}{\sqrt{2T}}, \frac{\Omega_Z}{\sqrt{2T}}\right) \cdot \left(I_{XX}\Omega_X, I_{YY}\Omega_Y, I_{ZZ}\Omega_Z\right)}{|\mathbf{L}|} = \frac{\sqrt{2T}}{|\mathbf{L}|} \tag{2.152}$$

que es una constante por serlo la energía cinética y la magnitud del impulso angular, por lo que este plano tangente es invariante.

Esto nos permite visualizar el movimiento del cuerpo rígido. El elipsoide de inercia se moverá con el cuerpo pero siempre será tangente a un plano

invariante perpendicular al impulso angular que se encuentra a una distancia d del centro del elipsoide, que coincide con el punto fijo del cuerpo. Como el punto de tangencia está sobre el eje instantáneo de rotación su velocidad instantánea con respecto al plano es cero; deducimos por tanto que en su movimiento el elipsoide de inercia rueda sin resvalar sobre el plano invariante y con él se mueve el cuerpo rígido.

Así, dadas unas condiciones iniciales, tendremos una energía y un vector de impulso angular que son integrales de movimiento, esto nos permite encontrar el plano invariante; en el tiempo inicial el elipsoide de inercia será tangente a ese plano y al transcurrir el tiempo simplemente rotará rodando sin resbalar sobre el plano. Al transcurrir el tiempo el punto de contacto trazará una curva sobre el elipsoide que lleva el nombre de poloda[7], en nuestro caso en que no hay fuerzas externas la poloda es una curva cerrada. Naturalmente el punto de tangencia trazará también una curva sobre el plano invariante, ésta se conoce como la herpoloda, la que en general no será cerrada ya que se trata de una roseta. En el caso en que el cuerpo rígido tenga una simetría, esto es que dos de sus momentos principales de inercia sean iguales, el elipsoide será simétrico, las polodas serán círculos y las herpolodas siempre serán cerradas.

Hemos resuelto así el problema del cuerpo rígido con un punto fijo cuando no actúan fuerzas sobre él y lo hemos hecho sin definir un sistema de coordenadas que establezca su posición en el espacio; hemos utilizado directamente las relaciones vectoriales y las integrales de movimiento; nótese que para hacer un análisis cualitativo no tuvimos que resolver las ecuaciones de Euler, claro esto nos impide obtener el detalle numérico de la evolución temporal de las variables del sistema, pero si hemos obtenido una descripción rigurosa del movimiento. El caso en que si se aplica una torca sobre el sistema es mucho más complejo y dejaremos su estudio para el siguiente capítulo.

Podríamos pasar al estudio de otros sistemas mecánicos, sin embargo lo revelado hasta ahora por los ejemplos expuestos nos ha permitido contemplar las características básicas de lo que podemos llamar la formulación vectorial de la mecánica clásica. No continuaremos con desarrollos o ejemplos, ya que esto forma el contenido de los cursos de mecánica vectorial que constituyen un antecedente del presente texto.

Podemos resumir señalando que la formulación vectorial de la mecánica, esto es, la expresión vectorial directa de las leyes de Newton, contiene toda la estructura física de la teoría y es covariante, es decir, no depende del sis-

[7]En general se define la poloda como la curva que traza, al trascurrir el tiempo, la intersección del eje de instantáneo de rotación con el elipsoide de inercia.

tema de coordenadas que utilicemos; es más como formulación no requiere siquiera que definamos un sistema de coordenadas. Sin embargo en lo general la expresión vectorial no se puede aplicar directamente a una situación concreta y muchas veces para poder describir un sistema específico debemos utilizar coordenadas. Esto nos lleva a trasladar las ecuaciones vectoriales a ecuaciones diferenciales para las componentes de los vectores y en este proceso perdemos en cierta forma la covariancia, ya que el proceso no se hace de una manera invariante para todos los posibles sistemas de coordenadas. De hecho esto constituye una de las dificultades al trabajar los problemas mecánicos en forma vectorial e inspira la búsqueda de formulaciones en términos de coordenadas que sean covariantes; esto es, una en que las ecuaciones de movimiento queden, en lo general, expresadas en coordenadas pero en forma invariante ante los cambios de coordenadas. Lograr esto simplificaría muchos de los cálculos que hay que hacer para obtener las ecuaciones de movimiento de sistemas concretos, aunque es importante señalar que el contenido físico sería el mismo, esto es, las leyes de Newton quedarían expresadas en una forma matemática distinta, pero no se agregaría ningún elemento físico nuevo. La búsqueda de estas cuestiones nos lleva a lo que conocemos como mecánica analítica y a su desarrollo conceptual, formal y práctico. Esto a su vez nos permitirá avanzar enormemente en la profundidad con que podemos entender la física del movimiento mecánico, así como al desarrollo de métodos y procedimientos muy poderosos para la solución de problemas mecánicos concretos. El desarrollo de la mecánica analítica será pues una parte medular de este texto.

Capítulo 3

Formulación lagrangiana de la mecánica.

En el capítulo anterior concluimos la necesidad de formular la mecánica clásica en términos de coordenadas, pero en forma covariante. Esto significa encontrar expresiones generales para las ecuaciones de movimiento en coordenadas, y hacerlo de tal forma que estas expresiones generales sean las mismas para cualquier sistema de coordenadas que utilicemos. Naturalmente de estas expresiones generales obtendremos, al aplicarlas a un caso físico concreto y en un sistema de coordenadas específico, las ecuaciones diferenciales del movimiento expresadas en el sistema de coordenadas escogido; en otro sistema de coordenadas las ecuaciones de movimiento se verán distintas, pero las ecuaciones generales en coordenadas de las que derivamos los casos concretos, por simple substitución, deben ser siempre las mismas; a esto nos referimos por una formulación en coordenadas covariante. Para que una formulación cumpla con este requisito las ecuaciones generales de movimiento deben involucrar únicamente cantidades escalares, esto es, la expresión general de las leyes de movimiento debe hacerse en términos de escalares y no de cantidades vectoriales como ocurre en la formulación vectorial, la que, si bien como expresión vectorial de las leyes de movimiento es covariante, su expresión en coordenadas y componentes no resulta covariante.

En este capítulo encontraremos la forma de expresar las leyes del movimiento utilizando cantidades escalares, con ello tendremos una formulación en coordenadas y covariante.

3.1. Espacio de configuración, coordenadas generalizadas y ecuaciones de movimiento.

Comencemos estudiando un sistema mecánico totalmente general, el de una colección de n partículas en el espacio físico Newtoniano de tres dimensiones que numeraremos con el índice $i = 1, \cdots, n$. Cada partícula está actuada por una fuerza que depende de su posición, de la posición de las demás partículas y del tiempo. En principio cualquier sistema mecánico en el que las fuerzas no dependan de velocidades, puede verse así.

Introduzcamos un sistema cartesiano de coordenadas en el espacio físico de tres dimensiones (x, y, z), denotaremos el vector de posición de cada partícula por $\mathbf{r}_i$, sus componentes en la base inducida por las coordenadas serán (x_i, y_i, z_i). Por otro lado la fuerza sobre cada partícula será el vector $\mathbf{F_i}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r_n}, t)$ con componentes $F_{x_i}, F_{y_i}, F_{z_i}$. Como se trata de un sistema cartesiano de coordenadas, la expresión en coordenadas de las ecuaciones vectoriales de movimiento es inmediata

$$\begin{aligned} m_i\ddot{x}_i &= F_{x_i}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t), \\ m_i\ddot{y}_i &= F_{y_i}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t), \qquad i = 1, \cdots, n \\ m_i\ddot{z}_i &= F_{z_i}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t). \end{aligned} \tag{3.1}$$

En vez de entender (3.1) como n ecuaciones que describen el movimiento de cada una de las partículas, podemos verlas como las ecuaciones que describen el movimiento de un punto que representa la situación del sistema mecánico en un espacio de $3n$ dimensiones. Este espacio de $3n$ dimensiones será el producto cartesiano de n copias del espacio físico tridimensional, uno por cada partícula. Podemos usar como coordenadas en este espacio a las generadas por las que establecimos anteriormente en el espacio físico, llegando así a un sistema cartesiano de $3n$ dimensiones $(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n)$. A este espacio lo llamaremos espacio de configuración y destacamos que se establece una correspondencia biunívoca entre los puntos de este espacio y la situación o configuración geométrica del sistema mecánico.

En el espacio de configuración y en las coordenadas establecidas, las ecuaciones de movimiento son las expresada en (3.1); pero no tenemos por que limitarnos a usar esas coordenadas, de hecho lo que nos proponemos hacer es expresar las ecuaciones de movimiento en cualquier sistema de coordenadas. Introduzcamos para ello un nuevo sistema de coordenadas en este espacio de $3n$ dimensiones que denotaremos por $(q^1, q^2, \cdots, q^{3n})$.[1]

[1] Hemos enumerado las coordenadas usando índices arriba. La razón de hacer esto se

Entre éstas y las coordenadas anteriores deberá haber una transformación de coordenadas, esto es una colección de funciones continuas y derivables que establezcan una relación biunívoca entre ellas:

$$\begin{aligned} q^1 &= q^1(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t), \\ &\vdots \\ q^{3n} &= q^{3n}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t), \end{aligned} \tag{3.2}$$

e inversamente

$$\begin{aligned} x_i &= x_i(q^1, \cdots, q^{3n}, t), \\ y_i &= y_i(q^1, \cdots, q^{3n}, t), \qquad i = 1, \cdots, n, \\ z_i &= z_i(q^1, \cdots, q^{3n}, t). \end{aligned} \tag{3.3}$$

Nótese que hemos incluido la posibilidad de que la transformación de coordenadas dependa del tiempo, esto es, el nuevo sistema de coordenadas puede estar en movimiento. Vemos que las nuevas coordenadas mezclan las antiguas coordenadas de las partículas, lo que nos da una visión más clara de que estamos trabajando en un espacio nuevo y más rico en posibilidades que el del conjunto de n espacios tridimensionales originales. A un sistema de coordenadas como éste es costumbre llamarlo sistema de coordenadas generalizadas.

En el espacio físico cada una de las partículas tendrá una cierta velocidad. Para la partícula i esta velocidad será un vector $\mathbf{v}_i$ cuyas componentes, en la base correspondiente al sistema de coordenadas cartesianas (x, y, z), serán $(\dot{x}_i, \dot{y}_i, \dot{z}_i)$. En el espacio de configuración no tiene sentido físico la velocidad del punto representativo del sistema, pero tendremos en coordenadas generalizadas una representación de las velocidades físicas dada por

$$\begin{aligned} \dot{q}^1 &= \sum_{i=1}^{n} \left(\frac{\partial q^1}{\partial x_i}\dot{x}_i + \frac{\partial q^1}{\partial y_i}\dot{y}_i + \frac{\partial q^1}{\partial z_i}\dot{z}_i\right) + \frac{\partial q^1}{\partial t}, \\ &\vdots \\ \dot{q}^{3n} &= \sum_{i=1}^{n} \left(\frac{\partial q^{3n}}{\partial x_i}\dot{x}_i + \frac{\partial q^{3n}}{\partial y_i}\dot{y}_i + \frac{\partial q^{3n}}{\partial z_i}\dot{z}_i\right) + \frac{\partial q^{3n}}{\partial t}, \end{aligned} \tag{3.4}$$

verá con detalle cuando veamos la formulación hamiltoniana de la mecánica. Por lo pronto tómese esto como una simple notación. Es importante ver que los índices utilizados en las ecuaciones 3.1 se refieren a la numeración de las partículas, mientras que estos últimos se refieren a las dimensiones del espacio de configuración.

cantidades a las que llamaremos velocidades generalizadas. Naturalmente esta expresión puede ser invertida y las componentes cartesianas de las velocidades de las partículas en el espacio físico, $\dot{x}_i$, $\dot{y}_i$ y $\dot{z}_i$, quedarán expresadas, en términos de las coordenadas y las velocidades generalizadas, por

$$\begin{aligned}
\dot{x}_i &= \sum_{j=1}^{3n} \frac{\partial x_i}{\partial q^j}\dot{q}^j + \frac{\partial x_i}{\partial t}, \\
\dot{y}_i &= \sum_{j=1}^{3n} \frac{\partial y_i}{\partial q^j}\dot{q}^j + \frac{\partial x_i}{\partial t}, \qquad i = 1, \cdots, n, \\
\dot{z}_i &= \sum_{j=1}^{3n} \frac{\partial z_i}{\partial q^j}\dot{q}^j + \frac{\partial x_i}{\partial t}.
\end{aligned} \tag{3.5}$$

Notemos que las velocidades generalizadas no son, en lo general, las componentes de la derivada con respecto al tiempo del vector de posición del punto representativo del sistema mecánico en el espacio de $3n$ dimensiones. Es más las velocidades generalizadas no tienen, en lo general, dimensiones de velocidad, por ejemplo, si una de las q es un ángulo, la correspondiente velocidad $\dot{q}$ tendrá unidades de velocidad angular. En resumen la cantidad $\dot{q}^i$ no es la componente de un vector velocidad en el espacio de configuración, es, simplemente, la derivada con respecto al tiempo de la coordenada q^i[2]. Para tratar expresiones vectoriales en coordenadas generalizadas deberíamos tomar una base y considerar la variación de ésta, pero esto no es lo que pretendemos, no trataremos con vectores en el espacio de configuración. Pasemos pues a trabajar con cantidades escalares.

Una de las principales cantidades escalares en el contexto de la mecánica es la energía cinética, veamos como se expresa en coordenadas generalizadas.

En las coordenadas de cada partícula la energía cinética es

$$T = \sum_{i=1}^{n} \frac{1}{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \tag{3.6}$$

que es un escalar. Ahora haciendo uso de la ecuación (3.5) podemos substituir las componentes de la velocidad de cada partícula por sus expresiones en coordenadas generalizadas y obtener la expresión del escalar energía cinética

[2]En la fórmula **** del capítulo anterior encontramos la expresión del vector velocidad de una partícula en coordenadas esféricas, nótese que las componentes de este vector no son simplemente las derivadas respecto al tiempo de las coordenadas.

en coordenadas generalizadas

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i \left[\sum_{jk=1}^{3n} \left(\frac{\partial x_i}{\partial q^j}\frac{\partial x_i}{\partial q^k} + \frac{\partial y_i}{\partial q^j}\frac{\partial y_i}{\partial q^k} + \frac{\partial z_i}{\partial q^j}\frac{\partial z_i}{\partial q^k} \right) \dot{q}^j \dot{q}^k \right. \tag{3.7}$$

$$+2\sum_{k=1}^{3n} \left(\frac{\partial x_i}{\partial t}\frac{\partial x_i}{\partial q^k} + \frac{\partial y_i}{\partial t}\frac{\partial y_i}{\partial q^k} + \frac{\partial z_i}{\partial t}\frac{\partial z_i}{\partial q^k} \right) \dot{q}^k \tag{3.8}$$

$$\left. + \left(\frac{\partial x_i}{\partial t} \right)^2 + \left(\frac{\partial y_i}{\partial t} \right)^2 + \left(\frac{\partial z_i}{\partial t} \right)^2 \right] \tag{3.9}$$

$$= \frac{1}{2}\sum_{jk=1}^{3n} \left[\sum_{i=1}^{n} m_i \left(\frac{\partial x_i}{\partial q^j}\frac{\partial x_i}{\partial q^k} + \frac{\partial y_i}{\partial q^j}\frac{\partial y_i}{\partial q^k} + \frac{\partial z_i}{\partial q^j}\frac{\partial z_i}{\partial q^k} \right) \right] \dot{q}^j \dot{q}^k \tag{3.10}$$

$$+\sum_{k=1}^{3n} \left[\sum_{i=1}^{n} m_i \left(\frac{\partial x_i}{\partial t}\frac{\partial x_i}{\partial q^k} + \frac{\partial y_i}{\partial t}\frac{\partial y_i}{\partial q^k} + \frac{\partial z_i}{\partial t}\frac{\partial z_i}{\partial q^k} \right) \right] \dot{q}^k \tag{3.11}$$

$$+\frac{1}{2} \left[\sum_{i=1}^{n} m_i \left(\left(\frac{\partial x_i}{\partial t} \right)^2 + \left(\frac{\partial y_i}{\partial t} \right)^2 + \left(\frac{\partial z_i}{\partial t} \right)^2 \right) \right] \tag{3.12}$$

$$= \frac{1}{2}\sum_{jk=1}^{3n} \mathcal{I}_{jk}\dot{q}^j\dot{q}^k + \sum_{j=1}^{3n} \mathcal{J}_j \dot{q}^j + \frac{1}{2}\mathcal{K}. \tag{3.13}$$

Los coeficientes $\mathcal{I}_{jk}, \mathcal{J}_j$ y $\mathcal{K}$, quedan definidos por la ecuación anterior y dependen de las coordenadas generalizadas y del tiempo. Notamos que si el sistema generalizado de coordenadas no está en movimiento, no habrá dependencia temporal y $\mathcal{J} = \mathcal{K} = 0$, en ese caso la energía cinética será una forma cuadrática positiva definida en las velocidades generalizadas. Los coeficientes $\mathcal{I}_{jk}$ son simétricos, esto es $\mathcal{I}_{jk} = \mathcal{I}_{kj}$, y, en caso de tener un sitema ortogonal de coordenadas, todos aquellos con $j \neq k$ serán cero, estos coeficientes podemos contemplarlos como las componentes de un tensor al que llamaremos tensor de inercia.

En el espacio físico podemos observar que

$$p_{x_i} = \frac{\partial T}{\partial \dot{x}_i}, \tag{3.14}$$

$$p_{y_i} = \frac{\partial T}{\partial \dot{y}_i}, \tag{3.15}$$

$$p_{z_i} = \frac{\partial T}{\partial \dot{z}_i}, \tag{3.16}$$

son las componentes del vector impulso lineal de la partícula i a lo largo de cada una de las coordenadas. En el espacio de configuración podemos, por analogía, definir el impulso generalizado correspondiente a la coordenada generalizada q^j por medio de

$$p_j = \frac{\partial T}{\partial \dot{q}^j}. \tag{3.17}$$

Podemos observar que los impulsos generalizados no son, en general, impulsos lineales; pero si una coordenada generalizada representa un desplazamiento lineal, el impulso generalizado correspondiente y tendrá las unidades de impulso lineal; en caso de que la coordenada sea angular, el impulso generalizado correspondiente tendrá unidades de impulso angular; esto justifica el llamar a estas cantidades impulsos generalizados. Recordamos que no debemos confundirlos con las componentes de algún vector en el espacio de configuración.

En el espacio físico en adición a la posición, la velocidad y el impulso lineal de cada partícula, tenemos otra cantidad física de importancia: la fuerza que actúa sobre cada una de las partículas. De hecho con estas cantidades queda totalmente definido el sistema.

Para trasladar el concepto de fuerza del espacio físico al espacio de configuración, de forma análoga a lo que hemos hecho para la velocidad y el impulso y sin apartarnos de nuestro programa, debemos utilizar una cantidad escalar. El escalar que usaremos es el trabajo realizado contra las fuerzas al hacer un desplazamiento virtual de las partículas.

Tomemos el sistema de partículas en el espacio físico y pensemos en un instante del tiempo, como si éste se hubiera detenido, hagamos un desplazamiento infinitesimal de la posición de cada una de las partículas, al conjunto de estos desplazamientos lo llamaremos un desplazamiento virtual. Lo llamamos virtual porque los desplazamientos son arbitrarios, esto es, no son los desplazamientos que cada partícula tendría al actuar las fuerzas y correr el tiempo y porque los desplazamientos los tomaremos suficientemente pequeños como para que podamos considerar que la fuerza que actúa sobre cada partícula, que depende de las posiciones, no cambia.

El trabajo que se realiza al hacer un desplazamiento virtual del sistema de partículas involucra las fuerzas y los desplazamientos de cada partícula, que son cantidades vectoriales, pero es una cantidad escalar y, por tanto, puede calcularse en el espacio de configuración y en coordenadas generaliza-

das a partir de su expresión en el espacio físico y en coordenadas cartesianas aplicando directamente la transformación de coordenadas correspondiente.

Denotemos por δx_i, δy_i y δz_i a las componentes, en el sistema (x, y, z), de un desplazamiento virtual de cada partícula en el espacio físico, entonces el trabajo realizado para hacer el desplazamiento virtual será

$$\delta W = \sum_{i=1}^{n} \left(F_{x_i} \delta x_i + F_{y_i} \delta y_i + F_{z_i} \delta z_i \right). \tag{3.18}$$

Las componentes del desplazamiento de cada partícula en el espacio físico se expresan en coordenadas generalizadas por medio de

$$\begin{aligned} \delta x_i &= \sum_{j=1}^{3n} \frac{\partial x_i}{\partial q^j} \delta q^j, \\ \delta y_i &= \sum_{j=1}^{3n} \frac{\partial y_i}{\partial q^j} \delta q^j, \qquad i = 1, \cdots, n, \\ \delta z_i &= \sum_{j=1}^{3n} \frac{\partial z_i}{\partial q^j} \delta q^j, \end{aligned} \tag{3.19}$$

entonces el trabajo hecho por el desplazamiento virtual expresado en coordenadas generalizadas será

$$\begin{aligned} \delta W &= \sum_{i=1}^{n} \left[\sum_{j=1}^{3n} \left(F_{x_i} \frac{\partial x_i}{\partial q^j} \delta q^j + F_{y_i} \frac{\partial y_i}{\partial q^j} \delta q^j + F_{z_i} \frac{\partial z_i}{\partial q^j} \delta q^j \right) \right] \\ &= \sum_{j=1}^{3n} \left[\sum_{i=1}^{n} \left(F_{x_i} \frac{\partial x_i}{\partial q^j} + F_{y_i} \frac{\partial y_i}{\partial q^j} + F_{z_i} \frac{\partial z_i}{\partial q^j} \right) \right] \delta q^j, \end{aligned} \tag{3.20}$$

en donde simplemente debemos expresar las componentes de la fuerza en el espacio físico F_{x_i} en coordenadas generalizadas. Definamos Q_j por medio de

$$Q_j = \sum_{i=1}^{n} \left[F_{x_i} \frac{\partial x_i}{\partial q^j} + F_{y_i} \frac{\partial y_i}{\partial q^j} + F_{z_i} \frac{\partial z_i}{\partial q^j} \right] \tag{3.21}$$

y entonces

$$\delta W = \sum_{j=1}^{3n} Q_j \delta q^j. \tag{3.22}$$

A las cantidades Q_j, que son funciones de las coordenadas generalizadas y del tiempo, las llamaremos fuerzas generalizadas, ya que representan un papel similar al de las fuerzas físicas en el cálculo del trabajo realizado al efectuar un desplazamiento virtual, como puede verse al comparar las ecuaciones (3.18) y (3.22). Naturalmente las fuerzas generalizadas no serán en general fuerzas, ni serán componentes de ningún vector fuerza, ni tendrán unidades de fuerza.

La similitud entre fuerzas físicas y fuerzas generalizadas puede ser extendida más allá de lo que hasta aquí hemos mencionado, veamos esto.

Pensemos en el sistema mecánico en el espacio físico, tomemos una de las partículas, digamos la iésima. Ahora, dejando todas las demás partículas en sus posiciones originales, desplacemos virtualmente a esa partícula en alguna de las direcciones de los ejes coordenados, digamos la dirección x, por una cantidad δx_i, el trabajo realizado será $\delta W = F_{x_i}\delta x_i$; esto es, la componente de la fuerza física que actúa sobre una partícula en la dirección de una coordenada, es el coeficiente por el que debemos multiplicar el desplazamiento virtual de la partícula en esa dirección para obtener el trabajo realizado por dicho desplazamiento.

Pasemos ahora al espacio de configuración de $3n$ dimensiones y hagamos algo similar. Tomemos una de las coordenadas generalizadas, digamos la jotaésima, y, dejando todas las demás coordenadas fijas, hagamos un desplazamiento δq^j en la dirección de la coordenada q^j. Naturalmente esto se reflejará en el espacio físico en desplazamientos de muchas de las partículas, en general, de todas ellas, lo que constituye un desplazamiento virtual del sistema mecánico que conllevará un trabajo δW . Este trabajo es una cantidad escalar por lo que podrá calcularse de igual forma en el espacio físico o en el de configuración y en coordenadas generalizadas, la ecuación (3.22) nos indica como hacerlo en estas últimas, tendremos

$$\delta W = Q_j \delta q^j. \tag{3.23}$$

Esta ecuación nos dice que la fuerza generalizada Q_j, es el coeficiente por el que hay que multiplicar al desplazamiento δq^j para obtener el trabajo hecho por el desplazamiento virtual correspondiente al cambio δq^j. Esto es lo mismo que ocurre para las fuerzas físicas en el espacio físico y nos permite tomar este hecho como una definición de fuerza generalizada que sería análoga o similar a la de fuerza física, esto es, la fuerza generalizada Q_j correspondiente a la coordenada q^j, es el coeficiente por el que hay que multiplicar un desplazamiento virtual δq^j para obtener el trabajo virtual realizado al ejecutar dicho desplazamiento.

Veamos ahora el caso de un sistema conservativo, esto es aquél en que las fuerzas físicas que actúan sobre cada partícula pueden ser derivadas de un potencial $V(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n, z_n, t)$ por medio de

$$\begin{aligned} F_{x_i} &= -\frac{\partial V}{\partial x_i}, \\ F_{y_i} &= -\frac{\partial V}{\partial y_i}, \\ F_{z_i} &= -\frac{\partial V}{\partial z_i}. \end{aligned} \tag{3.24}$$

La función V es un escalar por lo que podemos expresarla directamente en coordenadas generalizadas, por tanto

$$\frac{\partial V}{\partial q^j} = \sum_{i=1}^{n} \left(\frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q^j} + \frac{\partial V}{\partial y_i}\frac{\partial y_i}{\partial q^j} + \frac{\partial V}{\partial z_i}\frac{\partial z_i}{\partial q^j} \right), \tag{3.25}$$

y usando (3.24)

$$\frac{\partial V}{\partial q^j} = -\sum_{i=1}^{n} \left(F_{x_i}\frac{\partial x_i}{\partial q^j} + F_{y_i}\frac{\partial y_i}{\partial q^j} + F_{z_i}\frac{\partial z_i}{\partial q^j} \right), \tag{3.26}$$

pero de la expresión (3.182) obtenemos

$$Q_j = -\frac{\partial V}{\partial q^j}. \tag{3.27}$$

Esto es, las fuerzas generalizadas son derivables de la misma función potencial y usando el mismo proceso de derivación parcial que en el caso de las componentes cartesianas de las fuerzas físicas que actúan sobre cada partícula.

Es interesante señalar que en el espacio físico la fórmula (3.24) únicamente se aplica cuando usamos un sistema cartesiano de coordenadas como el (x, y, z); esto se debe al hecho de que la fuerza es el gradiente de la función potencial y si se utiliza un sistema no cartesiano de coordenadas debemos tomar en consideración la variación de los vectores de la base asociada a las coordenadas en cada punto.

Por otro lado en el espacio de configuración la fórmula (3.27) siempre será válida independientemente del sistema de coordenadas generalizadas que usemos. Esto se debe a que, las fuerzas generalizadas en el espacio de configuración, no son las componentes de un vector de $3n$ dimensiones que

actúa, como una superfuerza $3n$-dimensional, sobre el punto que representa la configuración geométrica de las partículas en el espacio físico; son únicamente los coeficientes por los que hay que multiplicar δq^j para obtener el trabajo virtual.

En resumen, hemos establecido en el espacio de configuración los conceptos de coordenadas generalizadas, velocidades generalizadas, impulsos generalizados y fuerzas generalizadas y los hemos relacionado con los conceptos físicos básicos de posición, velocidad, impulso y fuerza en el espacio físico, para hacer todo esto hemos hecho uso únicamente de cantidades escalares. Si logramos, utilizando únicamente cantidades escalares y estas cantidades generalizadas, establecer ecuaciones que sean equivalentes a las ecuaciones vectoriales de Newton en el espacio físico, habremos logrado nuestro objetivo, esto es, establecer las ecuaciones de movimiento en coordenadas y en forma covariante o sea con expresiones invariantes ante las transformaciones de coordenadas.

Las ecuaciones de movimiento en el espacio físico y en coordenadas cartesianas (x, y, z) son

$$\begin{aligned} \frac{d}{dt}p_{x_i} &= F_{x_i}, \\ \frac{d}{dt}p_{y_i} &= F_{y_i}, \qquad i = 1, \cdots, n, \\ \frac{d}{dt}p_{z_i} &= F_{z_i}. \end{aligned} \tag{3.28}$$

A estas ecuaciones no les podemos aplicar directamente la transformación de coordenadas (3.2) para obtener las ecuaciones de movimiento en el espacio de configuración en coordenadas generalizadas, ya que que las cantidades que en ellas intervienen no son escalares, son componentes de vectores. Por otro lado la similitud entre los conceptos de fuerza e impulso generalizados con sus correspondientes de fuerza e impulso físicos nos orilla a calcular la derivada temporal del impulso generalizado, veamos esto.

De la ecuación (3.17), usando (3.16) y el carácter escalar de la energía cinética obtenemos

$$p_j = \frac{\partial T}{\partial \dot{q}^j} = \sum_{i=1}^{n} \left(p_{x_i} \frac{\partial \dot{x}_i}{\partial \dot{q}^j} + p_{y_i} \frac{\partial \dot{y}_i}{\partial \dot{q}^j} + p_{z_i} \frac{\partial \dot{z}_i}{\partial \dot{q}^j} \right), \qquad j = 1, \cdots, 3n. \tag{3.29}$$

Ahora bien de (3.5) y (3.2) obtenemos que

$$\begin{aligned}\frac{\partial \dot{x}_i}{\partial \dot{q}^j} &= \frac{\partial x_i}{\partial q^j},\\ \frac{\partial \dot{y}_i}{\partial \dot{q}^j} &= \frac{\partial y_i}{\partial q^j},\\ \frac{\partial \dot{z}_i}{\partial \dot{q}^j} &= \frac{\partial z_i}{\partial q^j}.\end{aligned} \tag{3.30}$$

y por tanto

$$p_j = \sum_{i=1}^{n}\left(p_{x_i}\frac{\partial x_i}{\partial q^j} + p_{y_i}\frac{\partial y_i}{\partial q^j} + p_{z_i}\frac{\partial z_i}{\partial q^j}\right), \qquad j = 1, \cdots, 3n. \tag{3.31}$$

Tomemos ahora la derivada respecto al tiempo

$$\begin{aligned}\frac{d}{dt}p_j &= \sum_{i=1}^{n}\left(\dot{p}_{x_i}\frac{\partial x_i}{\partial q^j} + \dot{p}_{y_i}\frac{\partial y_i}{\partial q^j} + \dot{p}_{z_i}\frac{\partial z_i}{\partial q^j}\right)\\ &\quad + \sum_{i=1}^{n} m_i\left(\dot{x}_i\frac{d}{dt}\left(\frac{\partial x_i}{\partial q^j}\right) + \dot{y}_i\frac{d}{dt}\left(\frac{\partial y_i}{\partial q^j}\right) + \dot{z}_i\frac{d}{dt}\left(\frac{\partial z_i}{\partial q^j}\right)\right) =\\ &\quad \sum_{i=1}^{n}\left(\dot{p}_{x_i}\frac{\partial x_i}{\partial q^j} + \dot{p}_{y_i}\frac{\partial y_i}{\partial q^j} + \dot{p}_{z_i}\frac{\partial z_i}{\partial q^j}\right)\\ &\quad + \sum_{i=1}^{n} m_i\left(\dot{x}_i\frac{\partial \dot{x}_i}{\partial q^j} + \dot{y}_i\frac{\partial \dot{y}_i}{\partial q^j} + \dot{z}_i\frac{\partial \dot{z}_i}{\partial q^j}\right) =\\ &\quad \sum_{i=1}^{n}\left(\dot{p}_{x_i}\frac{\partial x_i}{\partial q^j} + \dot{p}_{y_i}\frac{\partial y_i}{\partial q^j} + \dot{p}_{z_i}\frac{\partial z_i}{\partial q^j}\right)\\ &\quad + \frac{\partial}{\partial q^j}\left(\sum_{i=1}^{n}\frac{1}{2}m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)\right) =\\ &\quad \sum_{i=1}^{n}\left(\dot{p}_{x_i}\frac{\partial x_i}{\partial q^j} + \dot{p}_{y_i}\frac{\partial y_i}{\partial q^j} + \dot{p}_{z_i}\frac{\partial z_i}{\partial q^j}\right)\\ &\quad + \frac{\partial T}{\partial q^j}.\end{aligned} \tag{3.32}$$

Ahora usamos las ecuaciones (3.28), esto es las leyes de Newton, para obtener

$$\frac{d}{dt}p_j = \sum_{i=1}^{n}\left(F_{x_i}\frac{\partial x_i}{\partial q^j} + F_{y_i}\frac{\partial y_i}{\partial q^j} + F_{z_i}\frac{\partial z_i}{\partial q^j}\right) + \frac{\partial T}{\partial q^j}, \qquad j = 1, \cdots, 3n, \tag{3.33}$$

pero usando la definición dada en la ecuación (3.182) llegamos a

$$\frac{d}{dt}p_j = Q_j + \frac{\partial T}{\partial q^j}, \qquad j = 1, \cdots, 3n, \tag{3.34}$$

recordando que $p_j = \partial T/\partial \dot{q}^j$ obtenemos

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^j}\right) - \frac{\partial T}{\partial q^j} = Q_j, \qquad j = 1, \cdots, 3n. \tag{3.35}$$

Estas últimas ecuaciones son las que estábamos buscando, están formuladas en términos de las coordenadas, las velocidades y las fuerzas generalizadas e involucran únicamente a la energía cinética que es un escalar, por tanto son covariantes. Esto es, su forma es la misma en cualquier sistema de coordenadas generalizadas, claro la expresión concreta de la energía cinética cambiará con las coordenadas y también lo hará la expresión de las fuerzas generalizadas, pero la forma de las ecuaciones no cambia, es totalmente general.

Por otro lado hemos introducido las ecuaciones de Newton para obtenerlas y es fácil ver que, si usamos a las coordenadas cartesianas de las partículas en el espacio físico como coordenadas generalizadas, cosa que siempre podemos hacer, las ecuaciones (3.35) se reducen a las (3.28) o sea se recuperan las leyes de Newton, por tanto las ecuaciones (3.35) son matemáticamente equivalentes a las de Newton.

Esto quiere decir que las ecuaciones (3.35) nos proporcionan un conjunto de $3n$ ecuaciones diferenciales cuyas soluciones $q^j(t)$, una vez transformadas a las coordenadas adecuadas, serán soluciones $x_i(t)$, $y_i(t)$ y $z_i(t)$ de las ecuaciones de Newton y, de manera inversa, las soluciones $x_i(t)$, $y_i(t)$ y $z_i(t)$ de las ecuaciones de Newton, al aplicar la transformación de coordenadas, proporcionan funciones $q^j(t)$ que serán solución de las ecuaciones (3.35).

Cuando el sistema es conservativo la ecuación (3.27) nos proporciona una ventaja adicional, ya que se simplifica enormemente el cálculo de las fuerzas generalizadas. Esto se debe, nuevamente, a que el potencial es un escalar y por tanto únicamente tenemos que aplicar la transformación de coordenadas para obtenerlo a partir de conocerlo en cualquier otro sistema de coordenadas. Para estos sistemas conservativos las ecuaciones (3.35) pueden escribirse como

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^j}\right) - \frac{\partial T}{\partial q^j} = -\frac{\partial V}{\partial q^j}, \qquad j = 1, \cdots, 3n. \tag{3.36}$$

En aquellos casos en que la función potencial no dependa de las velocidades generalizadas, que entre los de interés físico son mayoría, podemos definir

una nueva función escalar de las coordenadas, las velocidades y el tiempo por medio de

$$L = T - V; \tag{3.37}$$

con esta definición las ecuaciones (3.36) quedan como

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) - \frac{\partial L}{\partial q^j} = 0, \qquad j = 1, \cdots, 3n. \tag{3.38}$$

A las ecuaciones (3.35), (3.36) o (3.38) se les conoce como las ecuaciones de Lagrange y a la función definida por (3.37) se le llama función lagrangiana o simplemente lagrangiana del sistema físico.

Las ecuaciones de Lagrange nos proporcionan las ecuaciones de movimiento a partir de conocer la expresión de la energía cinética y las fuerzas generalizadas; para los sistemas conservativos, bastará conocer la energía cinética y el potencial o sea la lagrangiana del sistema, podremos decir que toda la información física sobre el sistema mecánico está contenida en una cantidad escalar, la función lagrangiana. Las ecuaciones de Lagrange constituyen, por tanto, una formulación completa de la mecánica clásica en coordenadas y covariante.

Veamos un ejemplo ilustrativo del proceso que hemos seguido. Tomemos una partícula de masa m_1 que se mueve sobre un plano y otra de masa m_2 que se mueve sobre una recta perpendicular al plano en que se mueve la primera masa. En este caso hemos restringido el espacio físico a dos dimensiones para la primera partícula y a una dimensión para la segunda, esto lo hacemos para tener, en este ejemplo ilustrativo, un espacio de configuración de tres dimensiones que podemos visualizar fácilmente. Para la primera partícula construimos un sistema cartesiano de dos dimensiones (x_1, y_1) cuyo origen colocamos en el punto de intersección del plano con la recta. Para la segunda partícula nos basta con una coordenada, x_2, con origen coincidente con el anterior. Esto se ilustra en la figura 3.1.

Entre las dos partículas consideremos que se establece una fuerza gravitacional. La restricción al movimiento hace que la primera partícula solo responda a la proyección de esta fuerza sobre el plano, ya que cualquier componente perpendicular a éste se anula por la restricción; en el caso de la segunda únicamente actúa la componente a lo largo de la recta.

El espacio de configuración es el producto cartesiano del plano con la recta, por lo que es un espacio de tres dimensiones. Nótese que no es el espacio físico, en este caso podemos decir que el espacio físico es, o se limita, para la primera partícula al plano y para la segunda a la recta. Un punto en el espacio de configuración determina conjuntamente la posición de ambas

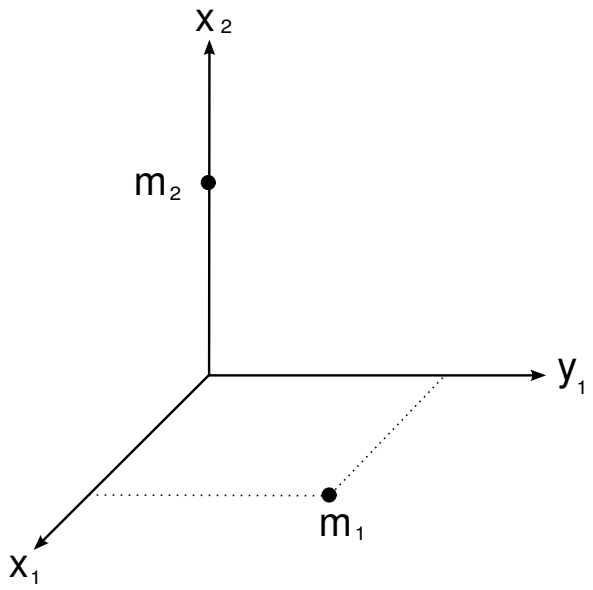


Figura 3.1: Espacio de configuración para dos partículas, una moviéndose en un plano y otra en una recta perpendicular a dicho plano.

partículas y no representa directamente a ninguna de ellas, ni es un punto material. Las coordenadas para el espacio de configuración que derivaríamos del producto cartesiano serían (x_1, y_1, x_2), estas serían las que en general hemos llamado x_i, y_i y z_i . Naturalmente que podemos usar otras coordenadas en el espacio de configuración, las que queramos, éstas serían las coordenadas generalizadas, (q^1, q^2, q^3); usemos como ejemplo coordenadas esféricas, como se ilustra en la figura 3.2, tendremos que $q^1 = r$, $q^2 = \theta$ y $q^3 = \varphi$.

Las ecuaciones de transformación entre los dos sistemas de coordenadas en el espacio de configuración son

$$\begin{aligned} r &= \sqrt{x_1^2 + y_1^2 + x_2^2}, \\ \theta &= \arctan \frac{\sqrt{x_1^2 + y_1^2}}{x_2}, \\ \varphi &= \arctan \frac{y_1}{x_1}, \end{aligned} \tag{3.39}$$

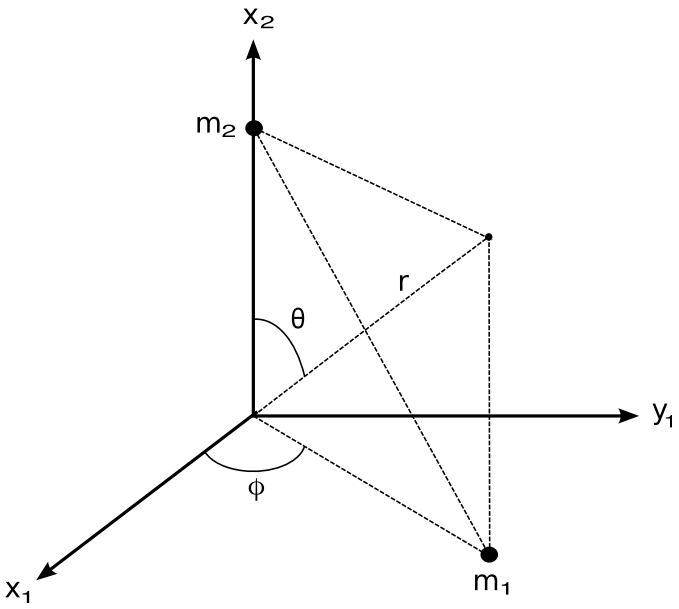


Figura 3.2: El espacio de configuración ilustrado en la figura 3.1 con coordenadas esféricas.

o a la inversa

$$\begin{aligned} x_1 &= r\text{sen}\theta\cos\varphi, \\ y_1 &= r\text{sen}\theta\text{sen}\varphi, \\ x_2 &= r\cos\theta. \end{aligned} \tag{3.40}$$

Estas ecuaciones son las ecuaciones (3.2) y (3.3) para este ejemplo.

Las ecuaciones (**??**) y (3.5) que relacionan las velocidades físicas con las velocidades generalizadas son

$$\begin{aligned} \dot{r} &= \frac{x_1\dot{x}_1 + y_1\dot{y}_1 + x_2\dot{x}_2}{\sqrt{x_1^2 + y_1^2 + x_2^2}}, \\ \dot{\theta} &= \frac{\dot{r}x_2 - r\dot{x}_2}{r\sqrt{x_1^2 + y_1^2}}, \\ \dot{\varphi} &= \frac{\dot{y}_1 x_1 - \dot{x}_1 y_1}{x_1^2 + y_1^2}; \end{aligned} \tag{3.41}$$

$$\begin{aligned} \dot{x}_1 &= \dot{r}\text{sen}\theta\cos\varphi + r\dot{\theta}\cos\theta\cos\varphi - r\dot{\varphi}\text{sen}\theta\text{sen}\varphi, \\ \dot{y}_1 &= \dot{r}\text{sen}\theta\text{sen}\varphi + r\dot{\theta}\cos\theta\text{sen}\varphi + r\dot{\varphi}\text{sen}\theta\cos\varphi, \\ \dot{x}_2 &= \dot{r}\cos\theta - r\dot{\theta}\text{sen}\theta. \end{aligned} \tag{3.42}$$

La energía cinética es un escalar por lo que podemos calcularla usando estas transformaciones y del hecho que en el espacio físico está dada por

$$T = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2\dot{x}_2^2, \tag{3.43}$$

obtenemos en coordenadas generalizadas que

$$\begin{aligned} T &= \frac{1}{2}m_1[(\dot{r}^2 + r^2\dot{\varphi}^2)\mathrm{sen}^2\theta + r^2\dot{\theta}^2\cos^2\theta + 2r\dot{r}\dot{\theta}\mathrm{sen}\theta\cos\theta] \\ &\quad + \frac{1}{2}m_2[\dot{r}^2\cos^2\theta + r^2\dot{\theta}^2\mathrm{sen}^2\theta - 2r\dot{r}\dot{\theta}\mathrm{sen}\theta\cos\theta], \end{aligned} \tag{3.44}$$

esto es, el tensor de inercia definido por (3.13) será

$$\mathcal{I}_{ij} = \begin{pmatrix} m_1\mathrm{sen}^2\theta + m_2\cos^2\theta & (m_1 - m_2)r\mathrm{sen}\theta\cos\theta & 0 \\ (m_1 - m_2)r\mathrm{sen}\theta\cos\theta & m_1r^2\cos^2\theta + m_2r^2\mathrm{sen}^2\theta & 0 \\ 0 & 0 & m_1r^2\mathrm{sen}^2\theta \end{pmatrix}. \tag{3.45}$$

Notemos que ésta no es la expresión de la energía cinética de una partícula en coordenadas esféricas, que sería

$$T_{\mathrm{esf}} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2\mathrm{sen}^2\theta). \tag{3.46}$$

Esto nos hace ver con claridad que el espacio de configuración no es el espacio físico.

Pasemos a analizar las fuerzas. En el espacio físico hemos dicho que cada partícula atrae a la otra con una interacción gravitacional, pero como hemos establecido que la primera partícula se mueve en un plano y la segunda en una recta, únicamente debemos tomar en cuenta las componentes de la fuerza sobre el plano para la primera partícula y la componente sobre la recta para la segunda. Podemos ver esto como si existieran unas fuerzas normales al plano y a la recta que obligan a las partículas a permanecer en ellos. La intensidad de la fuerza será m_1m_2G/d^2 donde $d = \sqrt{x_1^2 + y_1^2 + x_2^2}$.

Es entonces simple calcular la fuerza sobre cada partícula en las coordenadas cartesianas (x_1, y_1) y x_2 que hemos establecido para cada una de las partículas, obtendremos para la primera partícula

$$\begin{aligned} F_{x_1} &= -m_1m_2G\frac{x_1}{d^3}, \\ F_{y_1} &= -m_1m_2G\frac{y_1}{d^3}, \end{aligned} \tag{3.47}$$

y para la segunda

$$F_{x_2} = -m_1m_2G\frac{x_2}{d^3}. \tag{3.48}$$

Notemos que estas dos fuerzas son derivables de un potencial, ambas del mismo potencial, que es

$$V(x_1, y_1, x_2) = -\frac{m_1 m_2 G}{\sqrt{x_1^2 + y_1^2 + x_2^2}}. \tag{3.49}$$

Esto lo debemos entender por separado para cada partícula, esto es la primera partícula está sobre un plano por lo que el gradiente en ese plano y en las coordenadas que estamos usando es $(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1})$ y para la partícula dos que se mueve sobre una recta será simplemente $\frac{\partial}{\partial x_2}$.

Ahora podemos expresar la función potencial, que es un escalar, en las coordenadas generalizadas y veremos fácilmente que

$$V(r, \theta, \varphi) = -\frac{m_1 m_2 G}{r}, \tag{3.50}$$

que es independiente de θ y φ.

A partir de este potencial y de las ecuaciones (3.27) podemos calcular las fuerzas generalizadas y obtendremos

$$\begin{aligned} Q_r &= -\frac{m_1 m_2 G}{r^2}, \\ Q_\theta &= 0, \\ Q_\varphi &= 0. \end{aligned} \tag{3.51}$$

Podemos también calcular estas fuerzas generalizadas a partir de su definición usando desplazamientos virtuales, veamos esto.

Hagamos un pequeño desplazamiento $\delta\varphi$ dejando las otras dos coordenadas generalizadas fijas. En el espacio físico esto se traduce en (ver figura 3.2) dejar la partícula dos fija y desplazar la uno a lo largo de un pequeño arco centrado en la intersección del plano con la recta; la primera partícula no se desplaza por lo que no se realiza trabajo, en el caso de segunda la fuerza apunta al centro del círculo por el que se desplaza, por lo que la fuerza y el desplazamiento son perpendiculares y el trabajo será cero, de aquí que la fuerza generalizada también lo sea.

Hagamos ahora un desplazamiento $\delta\theta$ dejando las otras coordenadas generalizadas fijas, podemos ver que esto produce un desplazamiento virtual que involucra a las dos partículas. La primera se desplazará en dirección radial en su plano por una cantidad $r\delta\theta\cos\theta$ y la segunda lo hará a lo largo de su recta, por una cantidad $-r\delta\theta\sin\theta$. Esto conduce a los siguientes

desplazamientos en las direcciones de las coordenadas en el espacio físico

$$\begin{aligned} \delta x_1 &= \frac{x_2 x_1}{\sqrt{x_1^2+y_1^2}}\,\delta\theta, \\ \delta y_1 &= \frac{x_2 y_1}{\sqrt{x_1^2+y_1^2}}\,\delta\theta, \\ \delta x_2 &= -\sqrt{x_1^2+y_1^2}\,\delta\theta. \end{aligned} \tag{3.52}$$

Las fuerzas físicas están dadas por la ecuación (3.47), vemos de inmediato que el desplazamiento de la partícula dos va en contra de la fuerza que la actúa, mientras que el de la partícula uno va a su favor, por lo que los trabajos realizados tendrán signos opuestos y al hacer el cálculo encontramos que el trabajo total es cero, razón por la que Q_θ es cero.

Ahora para un desplazamiento en la dirección de r tendremos

$$\begin{aligned} \delta x_1 &= \frac{x_1}{d}\delta r, \\ \delta y_1 &= \frac{y_1}{d}\delta r, \\ \delta x_2 &= \frac{x_2}{d}\delta r, \end{aligned} \tag{3.53}$$

y de (3.47) obtenemos que

$$\begin{aligned} W &= -\frac{m_1 m_2 G}{d^4}(x_1^2+y_1^2+x_2^2)\delta r \\ &= -\frac{m_1 m_2 G}{r^2}\delta r. \end{aligned} \tag{3.54}$$

Esto implica que

$$Q_r = -\frac{m_1 m_2 G}{r^2}. \tag{3.55}$$

Vemos que el cálculo de las fuerzas generalizadas a partir del potencial o del trabajo hecho por los desplazamientos virtuales da los mismos resultados, naturalmente es mucho más sencillo hacerlo a partir del potencial, pero esto solo se puede hacer para los sistemas conservativos. Pasemos ahora a establecer las ecuaciones de movimiento.

La función lagrangiana $L = T - V$ será

$$\begin{aligned} L &= \frac{1}{2}m_1[(\dot r^2 + r^2\dot\varphi^2)\mathrm{sen}^2\theta + r^2\dot\theta^2\cos^2\theta + 2r\dot r\dot\theta\mathrm{sen}\theta\cos\theta] \\ &\quad +\frac{1}{2}m_2[\dot r^2\cos^2\theta + r^2\dot\theta^2\mathrm{sen}^2\theta - 2r\dot r\dot\theta\mathrm{sen}\theta\cos\theta] + \frac{m_1 m_2 G}{r}. \end{aligned} \tag{3.56}$$

A partir de ésta y usando las ecuaciones de Lagrange (3.38) podemos obtener las ecuaciones de movimiento.

Podemos ver que la coordenada φ no aparece en la lagrangiana, por tanto la derivada parcial $\partial L/\partial\varphi$ será cero y la ecuación de Lagrange correspondiente a φ nos indica que

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\varphi}}\right) = 0, \tag{3.57}$$

o sea, el impulso generalizado p_φ es una integral de movimiento. Haciendo el cálculo respectivo tendremos

$$p_\varphi = m_1 r^2\dot{\varphi}\text{sen}^2\theta = cte. \tag{3.58}$$

Las ecuaciones para las otras dos coordenadas resultan engorrosas por lo que no las escribiremos aquí, sin embargo para el caso en que las dos masas sean iguales, $m_1 = m_2 = m$, las expresiones se simplifican notablemente.

La lagrangiana será

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2\text{sen}^2\theta) + \frac{m^2G}{r}, \tag{3.59}$$

que es la misma lagrangiana que corresponde al problema de Kepler en coordenadas esféricas, ya que el primer término es la energía cinética de una partícula de masa m en coordenadas esféricas y el segundo corresponde al potencial gravitacional. La ecuación (3.58) seguirá siendo válida y nos permite eliminar la variable $\dot{\varphi}$ de las ecuaciones de movimiento. Es más podemos poner como condición inicial que p_φ sea cero, lo que corresponde a dar una velocidad inicial a la partícula dos en la dirección del centro de coordenadas, como $\dot{\varphi}$ será permanentemente cero, podemos simplificar aun más la lagrangiana que quedará como

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{m^2G}{r}, \tag{3.60}$$

que es la lagrangiana del problema de Kepler en dos dimensiones. Las ecuaciones de movimiento obtenidas de las ecuaciones de Lagrange serán

$$\begin{aligned} m\ddot{r} - mr\dot{\theta}^2 - \frac{m^2G}{r^2} &= 0, \\ \frac{d}{dt}\left(mr^2\dot{\theta}\right) &= 0, \end{aligned} \tag{3.61}$$

que son exactamente las ecuaciones que obtuvimos en el capítulo anterior para el sistema kepleriano. Nótese que al no aparecer en la lagrangiana la coordenada θ obtenemos la integral de movimiento $p_\theta = mr^2\dot{\theta}$.

Estos resultados indican que el punto representativo del sistema en el espacio de configuración trazará una órbita kepleriana, pero en realidad en el espacio físico la partícula uno se moverá sobre su plano y la partícula dos se moverá sobre su recta, estos dos movimientos se relacionan de tal suerte que, en el espacio de configuración, se obtiene una trayectoria kepleriana. Es interesante también ver que en caso de habernos planteado el problema de Kepler en dos dimensiones en coordenadas polares, hubiese sido inmediato escribir la lagrangiana (3.60) y de ella las ecuaciones de movimiento, esto nos indica la potencia del método lagrangiano.

Cuando las masas no son iguales la lagrangiana ya no es independiente de θ y tendremos una integral de movimiento menos; esto nos conduce a cuestiones similares a las comentadas para el péndulo doble en el capítulo anterior.

3.2. Constricciones.

Hemos visto como formular el problema mecánico general consistente en un conjunto de n partículas en interacción y sujetas a fuerzas externas en coordenadas y en forma covariante. La utilidad de esta formulación parece limitada por el hecho de que, si bien siempre es posible hacerlo, no siempre es conveniente estudiar los sistemas mecánicos como conjuntos de partículas. Tomemos, por ejemplo, el caso de un cuerpo rígido; podemos estudiarlo como el conjunto de moléculas que lo constituyen y considerar a cada una de ellas como una partícula sujeta a ciertas fuerzas. Esto nos obliga a tratar con un sistema de aproximadamente 10^{24} partículas e implica especificar la posición, la velocidad y la fuerza para cada una de ellas, esto es imposible y a todas luces inútil.

Por otro lado en un cuerpo rígido la posición de una de sus partículas está correlacionada con la de todas las demás, de tal suerte que la distancia entre cualesquiera dos permanece constante[3]. Tomemos una cualquiera de las 10^{24} partículas, su posición en el espacio se puede determinar por medio de tres coordenadas; tomemos cualquier otra partícula, como la distancia que la separa de la primera debe permanecer fija, solo podrá ocupar posiciones sobre una esfera centrada en la primera partícula, por lo que con dos parámetros podremos determinar su posición; al tomar una tercera partícula, que no esté

[3]De hecho, en la realidad hay pequeñísimos cambios dada la agitación molecular y la posible deformación elástica o plástica, pero para el estudio del movimiento mecánico es perfectamente valido considerarla constante, esto es pensar en un cuerpo total y perfectamente rígido.

sobre la linea que une a las otras dos, podrá moverse únicamente sobre un círculo cuyo centro se encuentra en esa linea, por tanto bastará un parámetro para fijar su posición. Naturalmente la posición de cualquier otra partícula queda determinada al fijar la de las tres partículas que hemos escogido, por tanto para especificar totalmente la situación en el espacio de un cuerpo rígido se requieren únicamente seis coordenadas y no 3×10^{24}.

Otro ejemplo lo constituye el péndulo esférico que estudiamos en el capítulo anterior. En este caso tenemos una partícula en el espacio físico de tres dimensiones, pero no tiene la capacidad de moverse libremente, ya que la varilla restringe su posición a la superficie de una esfera, sobre la cual bastan dos coordenadas para fijar su posición.

En muchísimos otros casos ocurre algo similar; los sistemas mecánicos, en general, consisten de partículas sujetas a restricciones en sus movimientos y estas restricciones reducen la dimensionalidad del sistema, esto es, el número mínimo de variables necesarias para determinar la situación del sistema en el espacio.

En el ejemplo del péndulo esférico pudimos resolver totalmente el problema mecánipermitióco utilizando únicamente dos coordenadas, que es la dimensionalidad real del sistema, ya que la restricción impuesta por la varilla reduce el movimiento a una superficie de dos dimensiones. Esto lo logramos hacer al descomponer las fuerzas que actúan sobre la partícula en dos componentes, una tangencial y otra normal a la esfera sobre la que se realiza el movimiento. La componente normal es la responsable de mantener a la partícula sobre la superficie de la esfera y, al ocurrir el movimiento natural, no realiza trabajo; esto nos permitió concentrarnos en la fuerza tangencial, única importante para el estudio de un movimiento que, por definición, está constreñido a permanecer sobre la esfera.

Esto se hace así en la formulación vectorial, debemos ver si es posible hacer algo similar en la formulación lagrangiana, desde luego, manteniendo la propiedad de covariancia. Para desarrollar esto debemos estudiar con mayor detenimiento las restricciones al movimiento.

En general llamaremos constricciones a las condiciones que restringen el movimiento de las partículas que componen un sistema mecánico. Estas constricciones las dividiremos en dos clases, que llamaremos holonómicas y no holonómicas.

Diremos que las constricciones son holonómicas cuando puedan establecerse como condiciones geométricas expresables, en términos de las coordenadas de cada partícula, por medio de ecuaciones de la forma

$$g_1(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t) = h_1, \tag{3.62}$$

$$\vdots \tag{3.63}$$

$$g_k(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t) = k_k, \tag{3.64}$$

en donde las h_i son constantes y las funciones g_i son diferenciables y linealmente independientes, esto es, su matriz Jacobiana no es singular. En caso contrario diremos que son no holonómicas.

Notamos que en el espacio de configuración estas ecuaciones definen un subespacio de dimensión $3n - k$ y que por ser diferenciables este subespacio será una variedad diferencial[4]. En el caso del péndulo esférico tenemos una sola partícula en tres dimensiones sujeta a una constricción dada, en coordenadas cartesianas, por la ecuación

$$x^2 + y^2 + z^2 = l^2, \tag{3.65}$$

por lo que tendremos una constricción holonómica, la variedad diferencial será la esfera. Otro ejemplo de constricciones holonómicas ocurre en el péndulo doble, en este caso tenemos dos partículas, cada una moviéndose en un espacio de dos dimensiones (x, y), el espacio de configuración es R^4 y en él podemos utilizar las coordenadas (x_1, y_1, x_2, y_2), las constricciones impuestas por las varillas se expresan por medio de

$$\begin{aligned} x_1^2 + y_1^2 &= l_1^2, \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 &= l_2^2. \end{aligned} \tag{3.66}$$

Aquí cabe un comentario ya que podría decirse que la matriz jacobiana en estos dos casos es singular cuando x_1 y y_1 o x_2 y y_2 son cero, sin embargo para alcanzar esto se requeriría que alguna de las l fuese cero, lo que implicaría que el péndulo doble se ha reducido a uno simple o a dos partículas constreñidas a permanecer fijas en un mismo punto.

Podemos tener constricciones holonómicas que no son estrictamente geométricas. Consideremos, por ejemplo, el caso de un disco plano que rueda sobre un riel (ver figura 3.3)

Geométricamente el disco es un cuerpo rígido de dos dimensiones, por tanto después de usar las constricciones que lo definen como tal, quedan únicamente tres dimensiones, por ejemplo la posición en el plano de su centro (x, y) y un ángulo θ de rotación. Ahora el disco ha de rodar sobre la recta x, esto hace que $y = r$, que es de la forma (3.64) y las dimensiones a considerar

[4]Más adelante estudiaremos el concepto de variedad diferencial, por lo pronto puede verse como una hipersuperficie inmersa en el espacio de configuración.

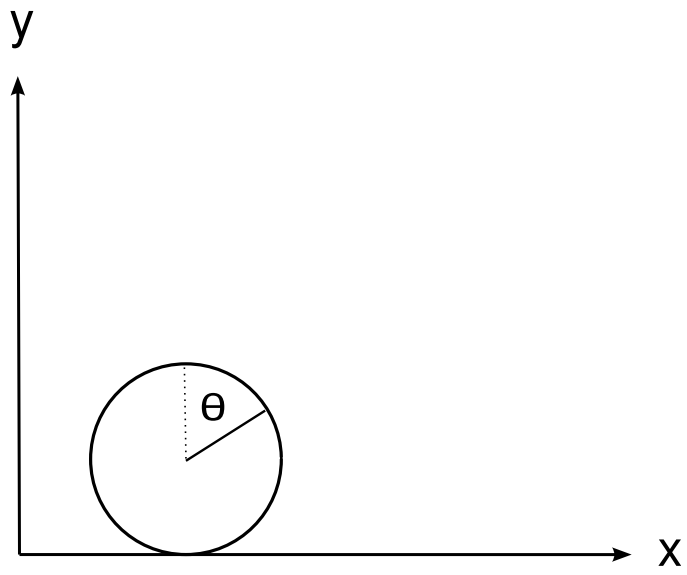


Figura 3.3: Un disco en el plano sujeto a la constricción de rodar sin resvalar sobre la recta x.

se han reducido a dos. Podemos ahora imponer una constricción no geométrica, que ruede sin resbalar, que es una condición cinemática. Rodar sin resbalar significa que la velocidad del punto de contacto del disco con el riel sea cero. Esta velocidad tendrá dos orígenes, el movimiento de translación del centro del disco y el producido por la rotación, la primera contribución es $\dot{x}$ y la segunda $r\dot{\theta}$ por lo que la condición de rodar sin resbalar se expresa por medio de

$$dx = rd\theta. \tag{3.67}$$

Ahora bien esta expresión es integrable y obtenemos de ella

$$x - r\theta = c, \tag{3.68}$$

donde c es una constante de integración. La condición de rodar sin resbalar se ha expresado con una ecuación de la forma (3.64), por tanto la constricción es holonómica.

En general al número de dimensiones que quedan después de considerar todas las constricciones lo llamamos el número de grados de libertad del sistema y coincide, en el caso holonómico, con la dimensión de la variedad diferencial inmersa en el espacio de configuración a la que se constriñe el movimiento; llamaremos a esta variedad la variedad de configuración.

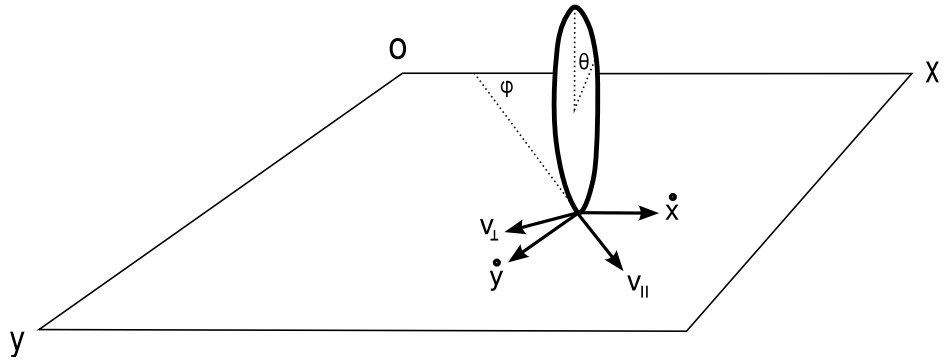


Figura 3.4: Un disco descansa verticalmente sobre un plano y esta constreñido a rodar sin resvalar.

No todas las constricciones resultan holonómicas. Por ejemplo consideremos el mismo disco del ejemplo anterior, pero ahora rodando sin resbalar sobre un plano con respecto al cual permanece siempre perpendicular. Su centro puede desplazarse en dos dimensiones (x, y) y para determinar su situación debemos proporcionar dos ángulos, uno de ellos, θ, nos indica el giro del disco sobre su propio eje y otro, φ, el giro del disco sobre un eje perpendicular al plano que pasa por su centro, ver la figura 3.4.

Tenemos cuatro coordenadas o dimensiones pero hemos considerado una constricción, que el disco ruede sin resbalar, o sea que la velocidad del punto de contacto del disco con el plano sea cero. Esta velocidad tendrá ahora dos componentes una tangencial al disco $v_{||}$ y otra perpendicular a este $v_\perp$. En términos de $\dot{x}$, $\dot{y}$, $\dot{\theta}$ y $\dot{\varphi}$ la primera vendrá dada por

$$v_{||} = \dot{x}\cos\varphi + \dot{y}\mathrm{sen}\varphi - r\dot{\theta}, \tag{3.69}$$

y la segunda por

$$v_\perp = -\dot{x}\mathrm{sen}\varphi + \dot{y}\cos\varphi. \tag{3.70}$$

Ambas componentes deben ser cero para que el disco ruede sin resbalar, por lo que las constricciones son

$$\begin{aligned} dx\cos\varphi + dy\mathrm{sen}\varphi - rd\theta &= 0, \\ -dx\mathrm{sen}\varphi + dy\cos\varphi &= 0. \end{aligned} \tag{3.71}$$

Ahora estas ecuaciones no son integrables; esto es, no existen dos funciones de las coordenadas cuyas diferenciales sean los miembros izquierdos de la ec. (3.71) multiplicados por algunos factores integradores. Una forma de ver

esto es notar que, de existir esas funciones, entonces dos de las coordenadas, digamos θ y φ quedarían determinas por las otras dos x y y; supongamos ahora que hemos comenzado el movimiento con el disco en el origen cuando los valores de θ y φ son cero, entonces si recorremos sin resbalar una curva sobre el plano que regrese al origen, deberíamos regresar con los mismos valores de θ y φ, sin embargo es fácil ver que podemos encontrar una curva que nos haga regresar con cualquier valor de esos ángulos, ya que el primero queda determinado por la longitud de la curva y el segundo por la tangente a la curva en su segundo paso por el origen. Concluimos que en este caso las constricciones no son holonómicas.

3.3. Mecánica lagrangiana de sistemas holonómicos.

Un sistema mecánico holonómico consiste de un conjunto de partículas sujetas a una serie de constricciones que se expresan por medio de ecuaciones de la forma (3.64). Al introducir en el espacio de configuración un sistema de coordenadas generalizadas siempre podremos utilizar las ecuaciones de constricción para definir k de esas coordenadas[5], esto es un sistema de coordenadas de la forma

$$\begin{aligned} q^1 &= g_1(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t), \\ &\vdots \\ q^k &= g_k(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t), \\ q^{k+1} &= q^{k+1}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t), \\ &\vdots \\ q^{3n} &= q^{3n}(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_n, y_n z_n, t), \end{aligned} \tag{3.72}$$

en donde hemos complementado a las k funciones g con $3n-k$ funciones diferenciables de tal forma que en conjunto las $3n$ funciones formen un sitema linealmente independiente, esto es, una transformación de coordenadas.

En estas coordenadas generalizadas las constricciones quedan expresadas por

$$q^i = h_i, \qquad i = 1, \cdots, k, \tag{3.73}$$

y las velocidades generalizadas correspondientes a estas k coordenadas serán cero. Por otra parte las fuerzas generalizadas correspondientes al resto de

[5]Esto lo podemos hacer porque las costricciones son holonómicas y, por definición, las funciones g son linealmente independientes y diferenciables.

las coordenadas, las que representan los grados de libertad del sistema, no dependerán de las coordenadas de constricción ya que, para cualquier desplazamiento virtual que respete las constricciones, las fuerzas de constricción no realizan trabajo alguno.

Esto implica que las ecuaciones de Lagrange del sistema pueden separarse en dos, las que corresponden a las primeras k coordenadas, las de constricción, y las correspondientes a los grados de libertad. Las primeras tendrán como solución a (3.73) que ya conocemos, por lo que no serán relevantes. Las segundas serán

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^j}\right) - \frac{\partial T}{\partial q^j} = Q_j, \qquad j = k+1, \cdots, 3n, \tag{3.74}$$

que resultan totalmente independientes de las fuerzas y coordenadas de constricción y, por tanto, suficientes para determinar el movimiento del sistema.

Desde luego si las fuerzas generalizadas son derivables de un potencial, podemos construir la función lagrangiana del sistema que quedará expresada en términos de las coordenadas correspondientes a los grados de libertad y de las constantes que definen las constricciones, en este caso las ecuaciones de movimiento serán

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) - \frac{\partial L}{\partial q^j} = 0, \qquad j = k+1, \cdots, 3n. \tag{3.75}$$

Podemos ver la gran potencia que se obtiene al aplicar una formulación en coordenadas y covariante a los sistemas holonómicos, podemos establecer las ecuaciones de movimiento únicamente para los grados de libertad del sistema y no considerar las constricciones, que se cumplirán automáticamente y, lo que es más importante, podemos ignorar las fuerzas de constricción, esto es, podemos establecer las ecuaciones de movimiento sobre la variedad de configuración sin conocer las fuerzas de constricción.

Desde el punto de vista del espacio de configuración, podemos decir que el movimiento ocurre sobre la variedad de configuración, esto es la varedad definida por las constricciones, y que en un sistema de 'coordenadas' generalizadas restringido a la variedad de configuración, las ecuaciones de movimiento son (3.74) o (3.75). Adicionalmente vemos que, sobre la variedad de configuración, con respecto a los cambios de coordenadas correspondientes a los grados de libertad, la formulación es covariante.[6]Hay que advertir que en

[6]Debemos recordar que para que esto funcione las coordenadas generalizadas para el espacio de configuración no pueden ser arbitrarias; deben tomarse tales que las ecuaciones de constricción se expresen como algunas de ellas iguales a una constante. El resto de las coordenadas pueden ser arbitrarias, claro consistentes con la condición de que entre todas formen un sistema de coordenadas generalizadas completo.

pleno rigor no se puede hablar de un sistema de coordenadas sobre la variedad de configuración ya que ésta no es en general homeomorfa a un espacio R^n.

Tomemos como ejemplo nuevamente el caso del péndulo esférico. Utilicemos en el espacio de tres dimensiones coordenadas esféricas, una de ellas es el radio r que cumple con los requerimientos para especificar la constricción, que vendría a quedar como $r = h$. Las otras dos θ y φ serán un sistema de coordenadas generalizadas para el problema, pero ello implicaría que θ y φ fueran coordenadas sobre la esfera y en rigor no se puede establecer un sistema de coordenadas sobre una esfera, veamos esto con detenimiento.

La transformación de coordenadas esféricas a cartesianas es

$$\begin{aligned} x &= r\text{sen}\theta\cos\varphi, \\ y &= r\text{sen}\theta\text{sen}\varphi, \\ z &= r\cos\theta. \end{aligned} \tag{3.76}$$

El determinante de la matriz jacobiana de esta transformación es $r^2\text{sen}\theta$; podemos ver que vale cero en $r = 0$, pero este valor de r no es alcanzable ya que la constricción es $r = h$ y si h fuera cero no habría péndulo ni movimiento; pero también es cero en $\theta = 0$ y en $\theta = \pi$, esto es en los polos de la esfera. Podemos ver que en esos dos puntos la coordenada φ es irrelevante, esto es, puede tomar cualquier valor, en esos dos puntos la trasformación no es biunívoca. Algo similar ocurre con cualquier otro intento de construir un sistema de coordenadas sobre una esfera; sin embargo podemos quitar esos dos puntos y pensar que nuestro sistema de coordenadas tiene un carácter local, esto es, que es valido para toda la esfera excepto para esos dos puntos.

Así cuando hablamos de coordenadas generalizadas sobre la variedad de configuración, deberíamos referirnos a un sistema de coordenadas locales. Ahora, en general, basta con quitar algunos puntos o algunas curvas, o sea la localidad puede abarcar casi toda la variedad y, en donde no es valido el sistema de coordenadas, por ejemplo los polos de la esfera en el péndulo esférico, podemos recurrir con facilidad a estudiar lo que le ocurre a las soluciones que pasan por esos puntos. En un terreno más riguroso la teoría de las variedades diferenciales resuelve totalmente estas cuestiones y nos enseña que, considerando a las coordenadas como locales y ampliando o extrapolando los resultados a los puntos en donde aparecen dificultades no se cometen errores. Entonces cuando nos referimos a un sistema de coordenadas generalizadas para la variedad de configuración, debemos entender que se trata de un sistema local de coordenadas, sin embargo para abreviar el lenguaje simplemente diremos coordenadas generalizadas, el lector entenderá que se

trata de un sistema local.

Pasemos a estudiar algunos ejemplos de aplicación de la formulación lagrangiana.

Como primer ejemplo consideremos nuevamente el péndulo esférico. Consiste, como hemos visto, en una partícula de masa m en el espacio de tres dimensiones sujeta a la constricción de permanecer a una distancia h de un punto fijo. Usando un sistema de coordenadas esféricas con origen en el punto fijo, la constricción se expresa por $r = h$, por lo que es holonómica. La variedad de configuración es la esfera y podemos utilizar los ángulos θ y φ como coordenadas generalizadas.

La energía cinética se puede calcular utilizando la transformación a coordenadas cartesianas con $r = h$. De ellas obtenemos

$$\begin{aligned} \dot{x} &= h\dot{\theta}\cos\theta\cos\varphi - h\dot{\varphi}\text{sen}\theta\text{sen}\varphi, \\ \dot{y} &= h\dot{\theta}\cos\theta\text{sen}\varphi + h\dot{\varphi}\text{sen}\theta\cos\varphi, \\ \dot{z} &= -h\dot{\theta}\text{sen}\theta. \end{aligned} \tag{3.77}$$

Como la energía cinética es un escalar y en coordenadas cartesianas se expresa por $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$, podemos calcularla usando directamente las ecuaciones (3.77). Obtenemos

$$T = \frac{1}{2}mh^2(\dot{\theta}^2 + \dot{\varphi}^2\text{sen}^2\theta). \tag{3.78}$$

La energía potencial será

$$V = mgh\cos\theta, \tag{3.79}$$

de estas dos expresiones obtenemos

$$L = T - V = \frac{1}{2}mh^2(\dot{\theta}^2 + \dot{\varphi}^2\text{sen}^2\theta) - mgh\cos\theta. \tag{3.80}$$

Utilizando las ecuaciones de Lagrange obtenemos las ecuaciones de movimiento

$$\begin{aligned} \ddot{\theta} - (\dot{\varphi}^2\cos\theta + \frac{g}{h})\text{sen}\theta &= 0, \\ \frac{d}{dt}(mh^2\dot{\varphi}\text{sen}^2\theta) &= 0. \end{aligned} \tag{3.81}$$

Podemos ver que éstas son exactamente las ecuaciones de movimiento para el péndulo esférico que obtuvimos en el capítulo anterior utilizando métodos vectoriales.

Pasemos a analizar el ejemplo del péndulo doble que tratamos en el capítulo anterior. Como se ve en la figura 8.58 hemos escogido como coordenadas generalizadas a los ángulos θ_1 y θ_2 que hacen con la vertical las dos varillas. Podemos ver que son buenas coordenadas para las constricciones, ya que, complementándolas con la distancia de la primera partícula al punto de sustentación y con la distancia de la segunda partícula a la primera, obtendríamos un sistema de coordenadas generalizadas para el espacio de configuración, que en este caso es de cuatro dimensiones, en ese sistema las constricciones se expresan igualando estas dos últimas coordenadas a las longitudes de las varillas. Tendremos, como vimos en el capítulo anterior, un variedad de configuración que es el toro de dos dimensiones y las coordenadas generalizadas θ_1 y θ_2 serán un sistema local de coordenadas para esa variedad.

En la figura 8.58 hemos puesto un sistema de ejes cartesianos (x, y) que nos auxiliarán en el cálculo de la energía cinética. Para esto recordamos que la energía cinética es un escalar y que en las coordenadas cartesianas auxiliares está dada por

$$T = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2). \tag{3.82}$$

A su vez tenemos que

$$\begin{aligned} x_1 &= h_1 \text{sen}\theta_1, \\ y_1 &= -h_1 \cos\theta_1, \\ x_2 &= h_1 \text{sen}\theta_1 + h_2 \text{sen}\theta_2, \\ y_2 &= -h_1 \cos\theta_1 - h_2 \cos\theta_2, \end{aligned} \tag{3.83}$$

por tanto

$$\begin{aligned} \dot{x}_1 &= h_1\dot{\theta}_1 \cos\theta_1, \\ \dot{y}_1 &= h_1\dot{\theta}_1 \text{sen}\theta_1, \\ \dot{x}_2 &= h_1\dot{\theta}_1 \cos\theta_1 + h_2\dot{\theta}_2 \cos\theta_2, \\ \dot{y}_2 &= h_1\dot{\theta}_1 \text{sen}\theta_1 + h_2\dot{\theta}_2 \text{sen}\theta_2. \end{aligned} \tag{3.84}$$

Haciendo las substituciones correspondientes y simplificando obtenemos

$$T = \frac{1}{2}(m_1 + m_2)h_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2(h_2^2\dot{\theta}_2^2 + 2h_1h_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2)). \tag{3.85}$$

Por otra parte la función potencial será

$$V = -(m_1 + m_2)h_1 g \cos\theta_1 - m_2 h_2 g \cos\theta_2, \tag{3.86}$$

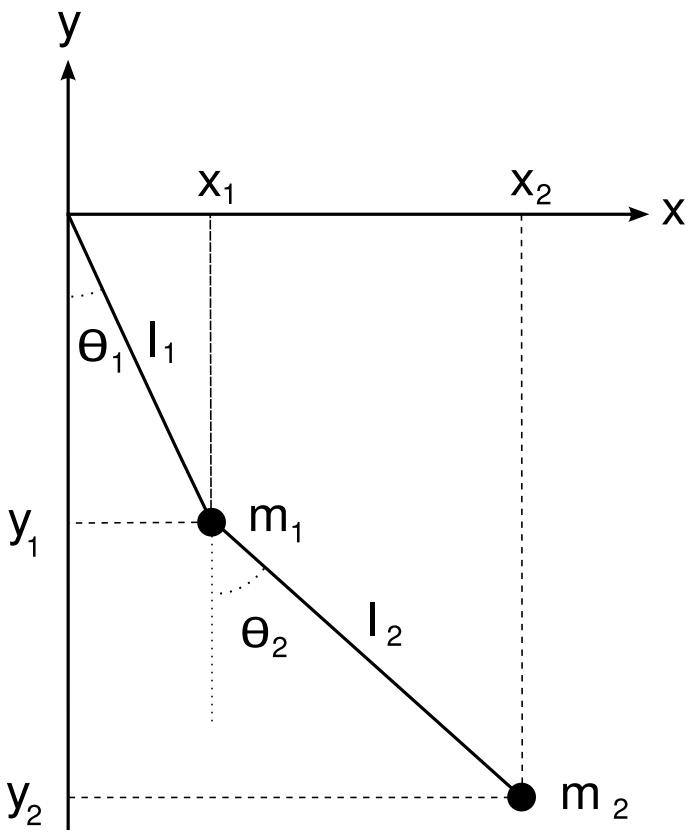


Figura 3.5: El péndulo doble. Las coordenadas θ_1 y θ_2 se utilizan como coordenadas generalizadas. Las coordenadas (x_1, x_2) y (y_1, y_2) son coordenadas cartesianas en el espacio físico.

por lo que la lagrangiana será

$$\begin{aligned} L = T - V = \; & \tfrac{1}{2}(m_1+m_2)h_1^2\dot{\theta}_1^2 + \tfrac{1}{2}m_2(h_2^2\dot{\theta}_2^2 + 2h_1h_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2)) \\ & +(m_1+m_2)h_1 g\cos\theta_1 + m_2h_2g\cos\theta_2. \end{aligned} \tag{3.87}$$

Para encontrar las ecuaciones de movimiento usamos las ecuaciones de Lagrange (3.38) . Hacemos el cálculo

$$\begin{aligned} \frac{\partial L}{\partial\dot{\theta}_1} &= (m_1+m_2)h_1^2\dot{\theta}_1 + m_2h_1h_2\dot{\theta}_2\cos(\theta_1-\theta_2), \\ \frac{\partial L}{\partial\dot{\theta}_2} &= m_2h_2\dot{\theta}_2 + m_2h_1h_2\dot{\theta}_1\cos(\theta_1-\theta_2), \\ \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}_1}\right) &= (m_1+m_2)h_1^2\ddot{\theta}_1 + m_2h_1h_2\ddot{\theta}_2\cos(\theta_1-\theta_2) \\ &\quad -m_2h_1h_2(\dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_2^2)\mathrm{sen}(\theta_1-\theta_2), \\ \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}_2}\right) &= m_2h_2^2\ddot{\theta}_2^2 + m_2h_1h_2\ddot{\theta}_1\cos(\theta_1-\theta_2) \\ &\quad +m_2h_1h_2(\dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_1^2)\mathrm{sen}(\theta_1-\theta_2), \\ \frac{\partial L}{\partial\theta_1} &= -m_2h_1h_2\dot{\theta}_1\dot{\theta}_2\mathrm{sen}(\theta_1-\theta_2) - (m_1+m_2)h_1g\mathrm{sen}\theta_1, \\ \frac{\partial L}{\partial\theta_2} &= m_2h_1h_2\dot{\theta}_1\dot{\theta}_2\mathrm{sen}(\theta_1-\theta_2) - m_2h_2g\mathrm{sen}\theta_2, \end{aligned} \tag{3.88}$$

y finalmente las ecuaciones de movimiento son

$$\begin{aligned} (m_1+m_2)h_1^2\ddot{\theta}_1 + m_2h_1h_2\ddot{\theta}_2\cos(\theta_1-\theta_2) & \\ +m_2h_1h_2\dot{\theta}_2^2\mathrm{sen}(\theta_1-\theta_2) + (m_1+m_2)h_1g\mathrm{sen}\theta_1 &= 0, \\ m_2h_2^2\ddot{\theta}_2 + m_2h_1h_2\ddot{\theta}_1\cos(\theta_1-\theta_2) - m_2h_1h_2\dot{\theta}_1^2\mathrm{sen}(\theta_1-\theta_2) & \\ +m_2h_2g\mathrm{sen}\theta_2 &= 0. \end{aligned} \tag{3.89}$$

Este ejemplo nos revela el enorme poder de la formulación lagrangiana para encontrar las ecuaciones de movimiento; naturalmente una vez obtenidas las ecuaciones de movimiento el problema de encontrarles solución puede ser muy difícil, como ocurre en este caso.

El lector habrá observado que tanto en las ecuaciones de transformación a coordenadas generalizadas (3.2), como en las que establecen las constricciones (3.64), hemos permitido una dependencia en el tiempo, esto quiere decir que tanto el sistema de coordenadas generalizadas como las constricciones pueden variar con el tiempo y las ecuaciones de movimiento seguirán siendo

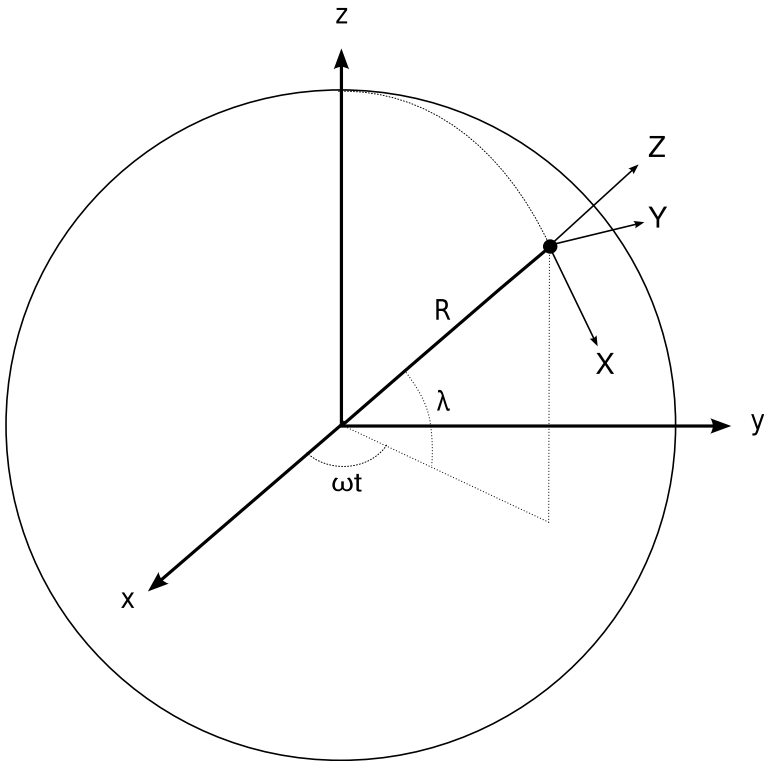


Figura 3.6: Coordenadas inerciales y fijas a la superficie de la tierra para el estudio del tiro parabólico.

las ecuaciones de Lagrange, la formulación toma en consideración estas posibles dependencias del tiempo. Esto es muy útil para tratar algunos sistemas en los que unas coordenadas en movimiento resultan ventajosas o aquellos en los que las constricciones varían con el tiempo.

Como ejemplo estudiaremos el tiro parabólico tomando en consideración la rotación de la Tierra. Para hacer esto consideremos un sistema de coordenadas inercial (x, y, z) con origen en el centro de la Tierra, ver figura 3.6.

Por otro lado consideremos un sistema de coordenadas generalizadas (X, Y, Z) fijo a la Tierra cuyo origen coincide con el punto en la superficie donde se realiza el tiro, su eje Z está a lo largo del radio terrestre que pasa por ese punto, su eje X está a lo largo del meridiano del lugar y apunta hacia el sur y el eje Y apunta hacia el este. Llamando R al radio de la Tierra, ω a su velocidad de rotación y λ a la latitud del lugar, tendremos, de la figura

3.6, que la transformación de coordenadas está dada por

$$\begin{aligned} x &= X\text{sen}\lambda\cos\omega t - Y\text{sen}\omega t + Z\cos\lambda\cos\omega t + R\cos\lambda\cos\omega t, \\ y &= X\text{sen}\lambda\text{sen}\omega t + Y\cos\omega t + Z\cos\lambda\text{sen}\omega t + R\cos\lambda\text{sen}\omega t, \\ z &= -X\cos\lambda + Z\text{sen}\lambda + R\text{sen}\lambda. \end{aligned} \tag{3.90}$$

La energía cinética en el sistema inercial será

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \tag{3.91}$$

haciendo la transformación correspondiente la energía cinética en las coordenadas generalizadas es

$$\begin{aligned} T &= \frac{1}{2}m\left[\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2 + 2\omega\left(X\dot{Y} - \dot{X}Y\right)\text{sen}\lambda\right. \\ &\quad +2\omega\left(R\dot{Y} + Z\dot{Y} - Y\dot{Z} + RX\omega\text{sen}\lambda + RZ\omega\cos\lambda + RX\omega\text{sen}\lambda\right) \qquad (3.92) \\ &\quad \left.+\,\omega^2\left(X^2\text{sen}^2\lambda + Y^2 + Z^2\cos^2\lambda + R^2\cos^2\lambda\right)\right]. \qquad (3.93) \end{aligned}$$

La fuerza de la gravedad la consideraremos como $-mg$ y en la dirección Z, esta aproximación es valida para distancias X y Y pequeñas comparadas con el radio de la Tierra. El potencial será entonces

$$V = mgZ. \tag{3.94}$$

De estas dos ecuaciones obtenemos la lagrangiana del sistema

$$L = \frac{1}{2}m\left[\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2 + 2\omega\left(X\dot{Y} - \dot{X}Y\right)\text{sen}\lambda + 2\omega\left(Z\dot{Y} - Y\dot{Z}\right)\cos\lambda\right. \tag{3.95}$$

$$\left.+\omega^2\left(2RX\text{sen}\lambda\cos\lambda + 2RZ\cos^2\lambda + 2XZ\text{sen}\lambda\cos\lambda + X^2\text{sen}^2\lambda + Y^2 + Z^2\cos^2\lambda\right)\right] \tag{3.96}$$

$$-mgZ. \tag{3.97}$$

Donde hemos eliminado los términos aditivos constantes y los proporcionales a una velocidad generalizada, ya que resultan irrelevantes para obtener las ecuaciones de movimiento.

De esta lagrangiana obtenemos las ecuaciones de movimiento

$$\begin{aligned} \ddot{X} &= 2\omega\dot{Y}\text{sen}\lambda + \omega^2(R+Z)\text{sen}\lambda\cos\lambda + \omega^2 X\text{sen}^2\lambda, \\ \ddot{Y} &= -2\omega(\dot{X}\text{sen}\lambda + \dot{Z}\cos\lambda) + \omega^2 Y, \\ \ddot{Z} &= 2\omega\dot{Y}\cos\lambda + \omega^2(R\cos\lambda + X\text{sen}\lambda)\cos\lambda + \omega^2 Z\cos^2\lambda - mg \end{aligned} \tag{3.98}$$

El lector podrá identificar en estas ecuaciones las aceleraciones centrífuga y de Coriolis.

Con estos ejemplos podemos ver la potencia del método lagrangiano y la forma en que la función lagrangiana resume las características físicas de un sistema mecánico.

Por otro lado la formulación lagrangiana nos permite desarrollar la teoría y estudiar propiedades físicas generales de los sistemas mecánicos, lo que profundiza nuestro conocimiento y nos permite contemplar a la mecánica desde perspectivas novedosas. Pasemos, pues, a ver algunas cuestiones en esta dirección.

3.4. Coordenadas ignorables.

De la forma misma de las ecuaciones de Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) - \frac{\partial L}{\partial q^j} = 0, \tag{3.99}$$

podemos ver que si la lagrangiana no depende explícitamente de alguna de las coordenadas generalizadas entonces la cantidad

$$\frac{\partial L}{\partial \dot{q}^j}, \tag{3.100}$$

es constante a lo largo de las soluciones del sistema, ya que su derivada total con respecto al tiempo es nula. Esto, en adición al sentido físico que la cantidad (3.100) pueda tener, nos permite reducir en una dimensión el problema de solucionar las ecuaciones de movimiento. Si en adición a esto la cantidad (3.100) no depende explicitamente del tiempo será una integral de movimiento.

Un ejemplo de esto ocurre en el caso del sistema kepleriano en coordenadas polares cuya lagrangiana está expresada en la ecuación (3.60); podemos ver que esta lagrangiana es independiente de la coordenada θ, la integral de movimiento correspondiente es el impulso angular $mr^2\dot{\theta}$.

En general, cuando ocurra que una de las coordenadas generalizadas q^i no aparece explícitamente en la lagrangiana de un sistema mecánico, diremos que esa coordenada es ignorable o cíclica. Recordando que a la expresión (3.100) la llamamos el impulso generalizado p_j, correspondiente a la coordenada generalizada q^j, tendremos que, si una coordenada es ignorable el impulso generalizado correspondiente es una constante de movimiento.

Es de observar que al expresar la misma lagrangiana en otro sistema de coordenadas puede resultar que no haya coordenadas ignorables, pero, desde

luego, se mantendrá el hecho de la existencia de la constante o integral de movimiento, simplemente su expresión será la que corresponda en el nuevo sistema de coordenadas y no será de la forma (3.100). Por ejemplo en coordenadas cartesianas la lagrangiana del sistema kepleriano es

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \frac{mMG}{\sqrt{x^2 + y^2}}, \tag{3.101}$$

vemos que ninguna de las dos coordenadas es ignorable, sin embargo el impulso angular, que en estas coordenadas se expresa por

$$m(x\dot{y} - y\dot{x}), \tag{3.102}$$

continuará siendo una integral de movimiento.

Un ejemplo de sumo interés, que ilustra la importancia de buscar sistemas de coordenadas en el que algunas son cíclicas, es el del trompo simétrico. Este es, un cuerpo rígido en que dos de sus momentos principales de inercia son iguales y que rota, en presencia de la gravedad, de tal forma que un punto de su eje de simetría se mantiene fijo. Usaremos la nomenclatura establecida en el capítulo anterior, esto es, tomamos un sitema inercial (x, y, z) cuyo origen coincide con el punto fijo y tomamos al eje z coincidente con la vertical. Construimos también un sistema de referencia (X, Y, Z) anclado al cuerpo con el mismo origen y colocado a lo largo de los ejes principales de inercia y ponemos el eje Z coincidente con el eje de simetría. Las componentes de los vectores con respecto a los ejes del sistema inercial las denotaremos con letras minúscula y con las correspondientes mayúsculas a las componentes con respecto al sistema anclado al cuerpo. Todo esto se muestra en la figura 3.7.

Para aplicar la formulación lagrangiana debemos encontrar unas coordenadas generalizadas que determinen la posición del cuerpo rígido. Podemos ver que determinar la posición del cuerpo y encontrar una transformación de coordenadas entre las del sistema inercial y las ancladas al cuerpo es equivalente. Ahora entre ambos sistema de coordenadas hay una rotación en tres dimensiones que puede representarse por una matriz de tres por tres, es decir por nueve valores, lo que es equivalente a un espacio R^9; naturalmente la matriz debe se ortogonal, lo que establece seis relaciones independientes entre sus elementos, por lo que únicamente se requieren tres parámetros para especificarla. Estas seis relaciones pueden verse como las ecuaciones que definen una variedad diferencial de dimensión tres inmersa en R^9, veamos de que variedad se trata. Una rotación general en tres dimensiones siempre puede verse como una rotación por un cierto ángulo en torno a un eje en

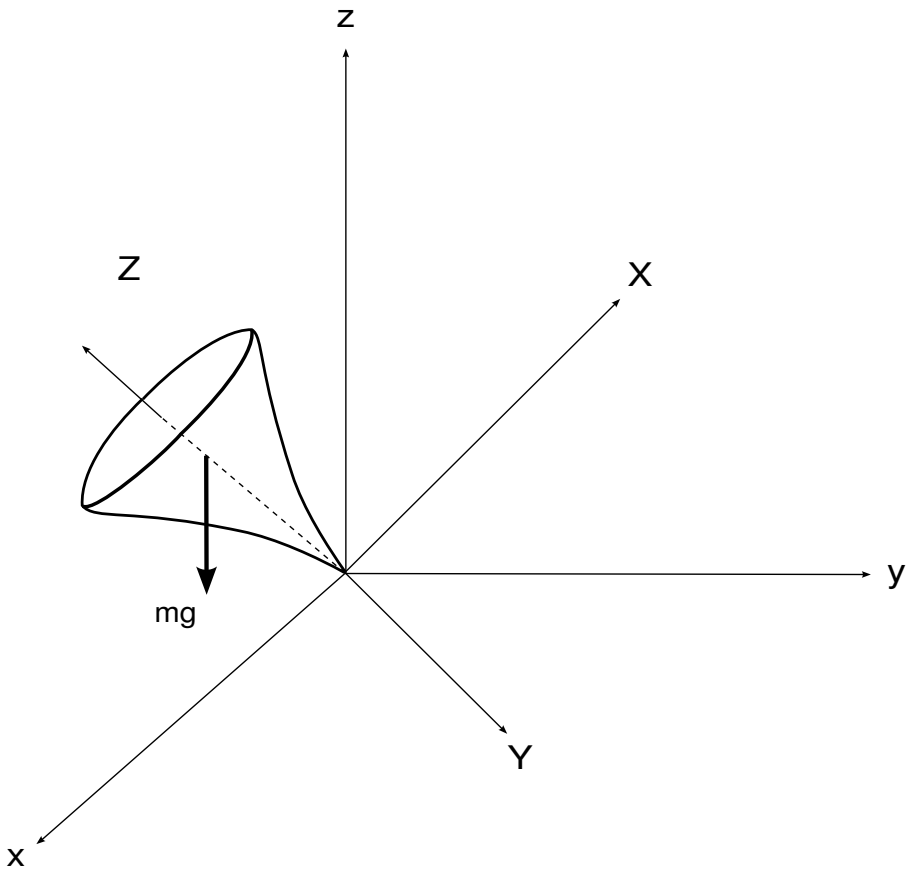


Figura 3.7:

cierta dirección. La dirección del eje puede identificarse con un punto sobre una esfera y el ángulo puede verse como un punto de un círculo, por ello la variedad será $S^2 \times S^1$. En realidad la variedad tiene dos componentes ajenos ya que el grupo ortogonal en tres dimensiones $O(3)$ incluye a las rotaciones puras y a las rotaciones compuestas con las reflexiones, estas últimas cambian la orientación y no son admisibles para el estudio de los movimientos de un cuerpo rígido, por ello debemos tomar únicamente una de las componentes, la correspondiente a las matrices ortogonales con determinante uno. Así el grupo de rotaciones en tres dimensiones, $SO(3)$, es homeomorfo a $S^2 \times S^1$, esto es, no es homeomorfo a R^3. Naturalmente no existen verdaderas coordenadas para una variedad como ésta, pero podemos encontrar coordenadas locales que cubren la mayor parte de la variedad, como ocurre en el caso del péndulo esférico.

Entre los posibles sistemas de coordenadas resultan particularmente útiles aquellos que tienen coordenadas cíclicas. El sistema presenta dos simetrías, una debida a que la fuerza de gravedad es siempre vertical, por lo que una rotación en torno a un eje vertical deja invariante al sistema y una segunda debida a la simetría propia de trompo; veremos más adelante que existe una relación entre las simetrías y las integrales de movimiento, razón por la cual buscamos un sistema de coordenadas que contenga a un ángulo de rotación en torno al eje vertical y otro en torno al eje del trompo, un sistema así es el integrado por los llamados ángulos de Euler.

En la figura 3.8 los ángulos de Euler quedan definidos como los ángulos (θ, ϕ, ψ). El ángulo θ es el formado entre los ejes z del sistema inercial y Z del anclado al cuerpo, esto hace que el plano (x, y) y el (X, Y) se intersecten en una recta que pasa por el origen a la que se llama linea de nodos y que hemos denotado con la letra n. El ángulo ϕ es el formado entre el eje x y la linea de nodos, vemos que se trata de una rotación en torno al eje vertical. Finalmente el ángulo ψ el formado entre la linea de nodos y el eje Y, representa una rotación en torno al eje del trompo. Es fácil ver de la figura que dada una posición del cuerpo tendremos una terna de valores de estos ángulos y que dada una terna de valores tendremos una posición, restringimos los valores de estos ángulos por

$$0 < \theta < \pi, \;\; 0 \leq \phi < 2\pi, \;\; 0 \leq \psi < 2\pi, \tag{3.103}$$

hemos excluido los valores cero y π de θ para evitarnos tener singularidades, naturalmente esto hace de los ángulos de Euler un sistema local de coordenadas.

En estas coordenadas el vector de velocidad angular puede descomponerse en tres componentes, una de magnitud $\dot{\theta}$ en la dirección de la linea

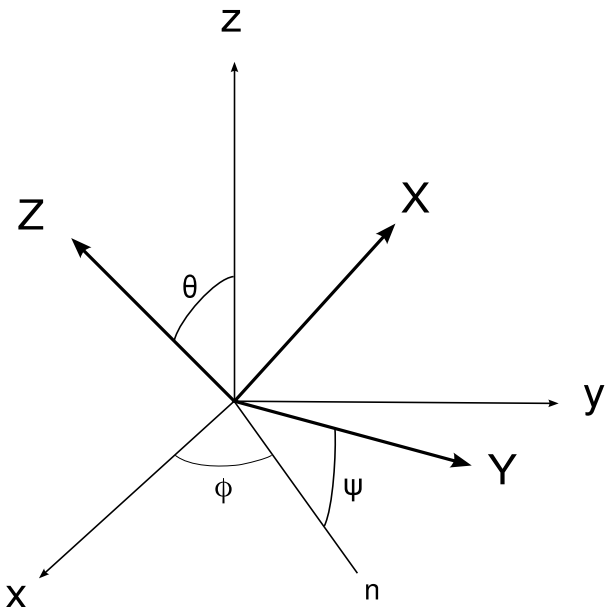


Figura 3.8: Los ángulos de Euler.

de nodos, una segunda de magnitud $\dot{\phi}$ en la dirección del eje z y una de magnitud $\dot{\psi}$ en la dirección del eje Z. De la figura 3.8 podemos ver que las componentes del vector velocidad angular a lo largo de los ejes del cuerpo pueden obtenerse de estas últimas por medio de

$$\Omega_X = \dot{\theta}\cos\psi + \dot{\phi}\text{sen}\theta\text{sen}\psi, \tag{3.104}$$

$$\Omega_Y = -\dot{\theta}\text{sen}\psi + \dot{\phi}\text{sen}\theta\cos\psi, \tag{3.105}$$

$$\Omega_Z = \dot{\psi} + \dot{\phi}\cos\theta. \tag{3.106}$$

La energía cinética puede calcularse fácilmente en las coordenadas del cuerpo, ya que en ellas el tensor de inercia es constante y diagonal, esto es

$$T = \frac{1}{2}\left(I_X\Omega_X^2 + I_Y\Omega_Y^2 + I_Z\Omega_Z^2\right), \tag{3.107}$$

o, en coordenadas (θ, ϕ, ψ)

$$T = \frac{1}{2}\left[I_X\left(\dot{\theta}\cos\psi + \dot{\phi}\text{sen}\theta\text{sen}\psi\right)^2 + \right. \tag{3.108}$$

$$I_Y\left(-\dot{\theta}\text{sen}\psi + \dot{\phi}\text{sen}\theta\cos\psi\right)^2 + \tag{3.109}$$

$$I_Z\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2\Big]. \tag{3.110}$$

La ecuación anterior proporciona la energía cinética para un cuerpo general; en el caso de un trompo simétrico $I_X = I_Y = I$, por lo que se reduce a

$$T=\frac{1}{2}\left[I\left(\dot{\theta}^2+\dot{\phi}^2\text{sen}^2\theta\right)+I_Z\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2\right]. \tag{3.111}$$

La única fuerza que interviene es la de la gravedad cuyo potencial es

$$V=mgh\cos\theta, \tag{3.112}$$

donde m es la masa total del trompo, h es la distancia entre el centro de masas y el punto fijo y g la aceleración de la gravedad. La lagrangiana será

$$L=\frac{1}{2}\left[I\left(\dot{\theta}^2+\dot{\phi}^2\text{sen}^2\theta\right)+I_Z\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2\right]-mgh\cos\theta, \tag{3.113}$$

y observamos que, en efecto, las coordenadas ϕ y ψ son cíclicas como buscábamos. La ecuación de Lagrange correspondiente a θ es

$$I\ddot{\theta}+\left(I_Z-I\right)\dot{\phi}^2\text{sen}\theta\cos\theta+I_Z\dot{\psi}\dot{\phi}\text{sen}\theta-mgh\text{sen}\theta=0, \tag{3.114}$$

y, como las coordenadas ϕ y ψ son cíclicas, p_ϕ y p_ψ serán integrales de movimiento; están dadas por

$$p_\phi=I\dot{\phi}\text{sen}^2\theta+I_Z\left(\dot{\psi}+\dot{\phi}\cos\theta\right)\cos\theta, \tag{3.115}$$

$$p_\psi=I_Z\left(\dot{\psi}+\dot{\phi}\cos\theta\right). \tag{3.116}$$

Podemos ver de (3.106) que

$$p_\psi=I_Z\Omega_Z, \tag{3.117}$$

por lo que la velocidad angular de giro en torno al eje de simetría del trompo es constante.

No requerimos resolver la ecuación en θ para encontrar el comportamiento del trompo ya que al substituir $\dot{\theta}$ y $\dot{\psi}$ por los valores que se obtienen para ellas de las integrales de movimiento, la energía total queda como

$$E=\frac{1}{2}I\dot{\theta}^2+\frac{1}{2I}\left[\frac{p_\phi-p_\psi\cos\theta}{\text{sen}\theta}\right]^2+\frac{1}{2}\frac{p_\psi^2}{I_Z}+mgh\cos\theta, \tag{3.118}$$

que podemos interpretar como la energía de una partícula ficticia sujeta a un potencial efectivo de la forma

$$V_{\text{ef}} = \frac{1}{2I}\left[\frac{p_\phi - p_\psi \cos\theta}{\text{sen}\theta}\right]^2 + mgh\cos\theta + \frac{1}{2}\frac{p_\psi^2}{I_Z}. \tag{3.119}$$

El tercer término, por tratarse de una constante, es en realidad intrascendente para el análisis de un potencial, pero lo incluimos ya que se trata de la energía cinética de rotación en torno del eje de simetría del trompo, que es constante.

Podemos ahora utilizar el método gráfico para estudiar el comportamiento del ángulo θ, para ello trazamos la gráfica de V_{ef} contra θ recordando que $0 < \theta < \pi$. Observamos que V_{ef} tiende a $+\infty$ tanto cuando θ tiende a cero y cuando tiende a π, por otro lado la derivada de V_{ef} con respecto a θ es

$$\frac{dV_{\text{ef}}}{d\theta} = \frac{1}{I}\frac{p_\phi p_\psi(1+\cos^2\theta) - (p_\phi^2 + p_\psi^2)\cos\theta}{\text{sen}^3\theta} - mgh\text{sen}\theta, \tag{3.120}$$

que únicamente tiene un cero para $0 < \theta < \pi$. Por tanto la gráfica de V_{ef} tiene la forma que se muestra en la figura 3.9.

Encontramos que θ oscila entre dos valores dados por las dos intersecciones, $\theta_{\text{mín}}$ y $\theta_{\text{máx}}$, de la gráfica de V_{ef} con el valor dado por las condiciones iniciales a la energía total; esto es, el eje de simetría del trompo presenta un movimiento periódico de cabeceo, que llamamos nutación, entre los valores $\theta_{\text{mín}}$ y $\theta_{\text{máx}}$. Como tenemos que

$$\dot{\phi} = \frac{p_\phi - p_\psi\cos\theta}{I\text{sen}^2\theta}, \tag{3.121}$$

a medida que ocurre la nutación el eje del trompo realizará un movimiento de rotación en torno a la vertical que llamamos precesión. Ahora bien, dependiendo de las condiciones iniciales, puede ocurrir que $\dot{\phi}$ tenga siempre el mismo signo o que éste cambie con los valores de θ a medida que ocurre la nutación; en el primer caso el eje del trompo trazará sobre una esfera centrada en el punto fijo una curva ilustrada en parte superior de la figura 3.10 y en el segundo la mostrada en la parte central; habrá un caso límite entre los dos en que trazará la curva mostrada en la parte inferior. Es interesante el hecho de que este último caso se obtiene si al tener un trompo girando sostenemos fijamente su eje de simetría a un cierto ángulo θ y lo soltamos, el trompo comenzará a nutar y a preceder, naturalmente lo hará en ese régimen intermedio. Recordamos que al tiempo que el trompo nuta y precede, gira a velocidad angular constante en torno a su eje, tendremos así

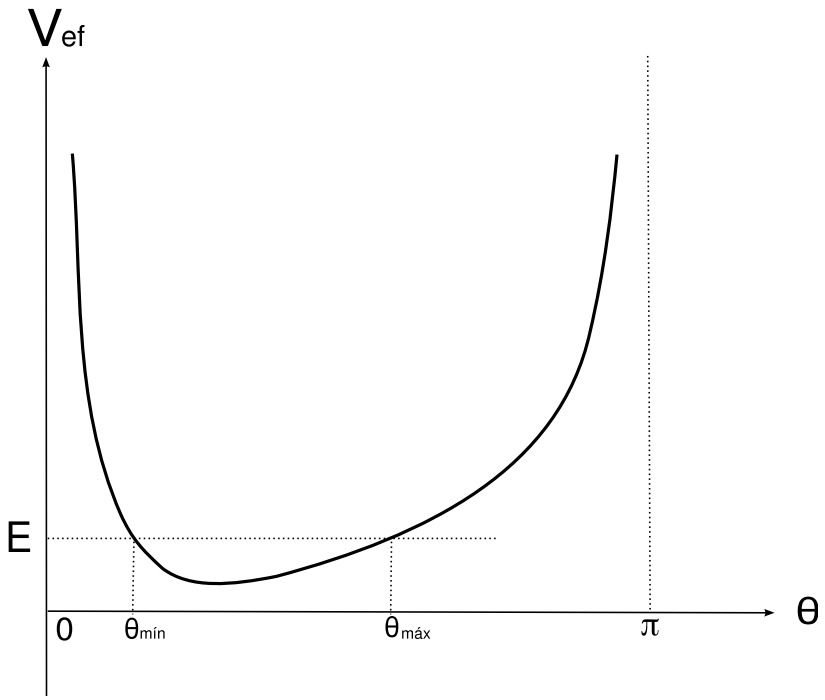


Figura 3.9: El potencial efectivo del trompo simétrico.

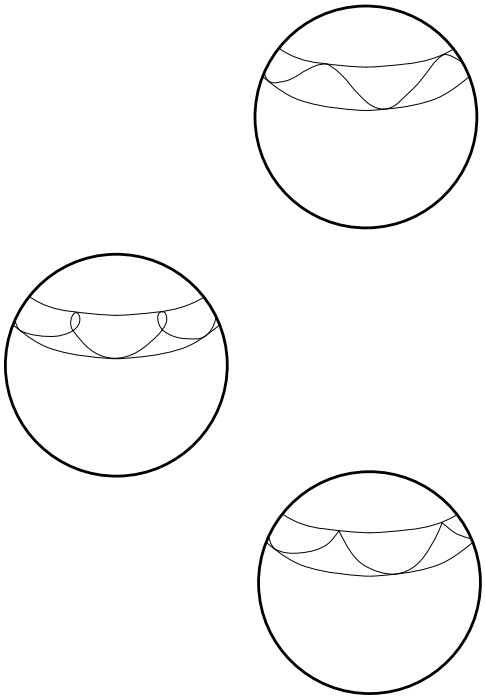


Figura 3.10: Curvas que traza el eje de un trompo simétrico sobre una esfera centrada en el punto fijo de su eje de simetría.

tres frecuencias, la de nutación, la de precesión y la de giro en torno del eje de simetría; si estas frecuencias son conmensurables entre si, el movimiento total será periódico, si son inconmensurables será cuasiperiódico, todo ello, claro está, determinado por las condiciones iniciales.

Un caso interesante ocurre cuando la energía coincide con el mínimo del potencial efectivo, entonces no habrá nutación, el movimiento de precesión será uniforme y el eje del trompo genera, en su movimiento, un cono circular.

Se dice que un trompo esta dormido cuando está en posición vertical y sin nutar. Este es un estado de equilibrio pues la torca producida por la gravedad es cero y el trompo está girando en torno a su eje de simetría. Sin embargo no siempre se trata de un equilibrio estable, esto es, aquel en que una pequeña perturbación genera una pequeña nutación, esto es, una oscilación entre dos ángulos θ muy cercanos a cero, veamos esto. Cuando θ es cero los ángulos de Euler no están bien definidos, por no estarlo la linea de los nodos, sin embargo el ángulo $\phi + \psi$ si lo está, de hecho es el ángulo de giro del trompo en torno a su eje de simetría. Las ecuaciones (3.116) y

(3.117) nos hacen ver que en este caso

$$p_\phi = p_\psi = I_Z\left(\dot{\phi} + \dot{\psi}\right) = I_Z\Omega_Z. \tag{3.122}$$

Para ángulos θ pequeños podemos aproximar $\cos\theta$ por $1 - \theta^2/2$ y $\sin\theta$ por θ, entonces el potencial efectivo se aproxima por

$$V_{ef} \approx \left(\frac{I_Z^2\Omega_Z^2}{8I} - \frac{1}{2}mgh\right)\theta^2 + mgh + \frac{1}{2}I_Z\Omega_Z^2. \tag{3.123}$$

En $\theta = 0$ tendremos entonces un mínimo si el coeficiente que multiplica a θ^2 es positivo y un máximo si es negativo; el mínimo corresponde al caso estable y el máximo al inestable, por tanto Ω_Z, la velocidad angular en torno al eje del trompo tendrá que ser, en valor absoluto, mayor que

$$\frac{2}{I_Z}\sqrt{Imgh} \tag{3.124}$$

para que el trompo quede dormido. Cuando se baila un trompo, al principio del movimiento está dormido, pero la fricción le hace perder energía y llega el momento en que Ω_Z se hace menor al valor crítico 3.124, en ese momento el trompo comienza a cabecear significativamente.

Hemos hecho un análisis del movimiento del trompo simétrico, naturalmente para obtener una solución cuantitativa completa debemos integrar la ecuación 3.118 y obtener $\theta(t)$, esto se puede hacer en términos de funciones elípticas, pero no lo haremos aquí por escapar al interés de este texto.

El caso de trompo asimétrico, esto es cuando $I_X \neq I_Y$, resulta mucho más complejo. Se pierde una de las integrales de movimiento p_ψ, aunque p_ϕ, la proyección del impulso angular total en la dirección vertical, continua siendo una integral de movimiento como podemos observar de la fórmula (3.110). La dificultad es similar a la relatada para el caso del péndulo doble, el sistema resulta no integrable y presenta trayectorias caóticas.

3.5. Integrales de movimiento para lagrangianas independientes del tiempo y el teorema del virial.

Pasemos ahora a analizar la expresión

$$\sum_{i=1}^{n} \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L, \tag{3.125}$$

que llamaremos cantidad de Jacobi. Calculemos su derivada total con respecto al tiempo

$$\frac{d}{dt}\left[\sum_{i=1}^{n}\dot{q}^i\frac{\partial L}{\partial \dot{q}^i}-L\right] \tag{3.126}$$

$$=\sum_{i=1}^{n}\left[\ddot{q}^i\frac{\partial L}{\partial \dot{q}^i}+\dot{q}^i\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)-\dot{q}^i\frac{\partial L}{\partial q^i}-\ddot{q}^i\frac{\partial L}{\partial \dot{q}^i}\right]-\frac{\partial L}{\partial t} \tag{3.127}$$

$$=\sum_{i=1}^{n}\dot{q}^i\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)-\frac{\partial L}{\partial q^i}\right]-\frac{\partial L}{\partial t}; \tag{3.128}$$

a lo largo de las trayectorias se deben cumplir las ecuaciones de Lagrange, por lo que los factores bajo el sumando son todos cero y entonces

$$\frac{d}{dt}\left[\sum_{i=1}^{n}\dot{q}^i\frac{\partial L}{\partial \dot{q}^i}-L\right]=-\frac{\partial L}{\partial t}. \tag{3.129}$$

Vemos que si la lagrangiana no depende explícitamente del tiempo la cantidad 3.125 es una integral de movimiento; esto es, todo sistema en el que la lagrangiana no dependa explícitamente del tiempo, tiene al menos una integral de movimiento, la cantidad de Jacobi.

En los casos en que la energía cinética sea una forma cuadrática homogénea de las velocidades generalizadas y el potencial no dependa de las velocidades, tendremos que, por el teorema de Euler para las funciones homogéneas,[7]

$$\sum_{i=1}^{n}\dot{q}^i\frac{\partial L}{\partial \dot{q}^i}=\sum_{i=1}^{n}\dot{q}^i\frac{\partial T}{\partial \dot{q}^i}=2T, \tag{3.130}$$

por tanto la expresión (3.125) será la energía total y para esos casos, la independencia temporal de la lagrangiana, implica la conservación de la energía total. En la expresión (3.97) para la lagrangiana del tiro parabólico podemos ver que la energía cinética no está expresada como una forma cuadrática, pero la lagrangiana es independiente del tiempo, deducimos que, si bien existe una integral de movimiento o cantidad conservada, ésta no es la energía total; mientras que la lagrangiana del péndulo doble (3.87) nos muestra que en ese caso la energía total se conserva.

[7] Una función $f(x_1, x_2, \cdots, x_n)$ es homogénea de grado k si $f(sx_1, sx_2, \cdots, sx_n) = s^k f(x_1, x_2, \cdots, x_n)$. El teorema de Euler nos dice que si $f(x_1, x_2, \cdots, x_n)$ es homogénea de grado k y diferenciable, entonces $\sum_i \frac{\partial f}{\partial x_i} x_i = kf$.

Pasando a un asunto distinto, la expresión (3.130) nos permite encontrar una relación entre el promedio de la energía cinética y el de la potencial, en lo que se conoce como el teorema del virial. Veamos esto.

Dada una función f de las coordenadas y las velocidades generalizadas su promedio temporal será

$$< f >= \lim_{\tau\to\infty} \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} f dt. \tag{3.131}$$

En los casos en que $f = dF/dt$ siendo F una función acotada

$$< f >= \lim_{\tau\to\infty} \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} \frac{dF}{dt} dt = \lim_{\tau\to\infty} \frac{F(\tau/2) - F(-\tau/2)}{\tau} = 0. \tag{3.132}$$

De (3.130) tenemos que

$$2T = \sum_{i=1}^{n} \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} = \frac{d}{dt} \left(\sum_{i=1}^{n} q^i \frac{\partial L}{\partial \dot{q}^i} \right) - \sum_{i=1}^{n} q^i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right). \tag{3.133}$$

Cuando tenemos un sistema mecánico y el movimiento esta acotado, esto es, cuando ninguna de las coordenadas ni de las velocidades puede hacerse infinita, el promedio del primer término del miembro derecho de esta ecuación será cero y obtendremos que

$$2 < T >= - < \sum_{i=1}^{n} q^i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) > . \tag{3.134}$$

Ahora como el promedio se toma a lo largo de una trayectoria del sistema, usando las ecuaciones de Lagrange, tendremos

$$2 < T >= - < \sum_{i=1}^{n} q^i \frac{\partial L}{\partial q^i} > \tag{3.135}$$

Si adicionalmente los coeficientes de la forma cuadrática que expresa la energía cinética son independientes de las coordenadas entonces

$$2 < T >=< \sum_{i=1}^{n} q^i \frac{\partial V}{\partial q^i} >, \tag{3.136}$$

De aquí vemos que si, en adición a lo demás, el potencial es una función homogénea de grado k de las coordenadas tendremos

$$2 < T >= k < V > . \tag{3.137}$$

Este resultado se conoce como el teorema del virial, en resumen establece que: En un sistema mecánico tal que: existe algún sistema de coordenadas en las que la energía cinética se expresa como una forma cuadrática de las velocidades con coeficientes constantes, las trayectorias permanecen acotadas y el potencial es una función homogénea de grado k, el promedio temporal de la energía cinética y el del potencial están relacionados por (3.137). Un ejemplo lo constituye el problema de Kepler para órbitas acotadas, esto es para energía total negativa, en este caso el potencial es una función homogénea de grado -1 y por tanto

$$< T >= -\frac{1}{2} < V > . \tag{3.138}$$

3.6. Las fuerzas de constricción.

Hemos visto que la formulación lagrangiana nos proporciona un método para encontrar las ecuaciones de movimiento de los sistemas holonómicos con el cual podemos hacer a un lado las fuerzas de constricción y trabajar únicamente con aquellas coordenadas que representan a los grados de libertad del sistema. Veamos ahora como esta formulación nos permite calcular las fuerzas de constricción, esto es, las fuerzas que los diversos elementos estructurales del sistema deben ejercer para mantener las constricciones durante el movimiento, por ejemplo la fuerza que debe proporcionar la varilla de un péndulo para mantener la masa sobre el círculo correspondiente. Este tipo de cálculos es importante por sí y en las aplicaciones, por ejemplo, cuando debemos conocer esas fuerzas para diseñar los elementos estructurales de una máquina.

En la ecuación (3.72) hicimos que las primeras coordenadas generalizadas estuvieran definidas por las ecuaciones de constricción, por tanto, en las ecuaciones de Lagrange correspondientes a esa coordenadas, las fuerzas generalizadas deben estar determinadas por las fuerzas necesarias para mantener las constricciones. Cuando tenemos un sistema con constricciones, podemos plantear la lagrangiana y las ecuaciones de Lagrange en forma general, esto es imaginando que las coordenadas correspondientes a las constricciones pudieran variar y que sus velocidades generalizadas no fuesen cero; claro, si las fuerzas generalizadas correspondientes a las coordenadas de constricción fuesen las que realmente ocurren en el sistema con constricciones, entonces estas coordenadas tomarían los valores constantes correspondientes y sus velocidades serían cero. Por otro lado sabemos que podemos resolver el problema con constricciones considerando únicamente las coordenadas y velocidades

generalizadas correspondientes a los grados de libertad, por tanto, si en los miembros izquierdos de las ecuaciones de Lagrange para las coordenadas correspondientes a las constricciones, substituimos las soluciones del sistema que obedece las constricciones, obtendremos las fuerzas generalizadas necesarias para mantener la constricción y de ellas podemos deducir las fuerzas físicas de constricción. Veamos algunos ejemplos.

Una maquina de Atwood consiste en una polea por la que pasa una cuerda de cuyos extremos cuelgan dos masas, ver la figura 3.11. Consideraremos que el movimiento de las masas es siempre vertical, que están sujetas a la acción de la gravedad y que la polea tiene masa despreciable. Como las masas solo se mueven en el sentido vertical se trata de un sistema de dos grados de libertad, la posición vertical de cada una de las partículas. Hay una constricción, el hecho de que la longitud de la cuerda, sin considerar la parte constante que queda en la polea, l sea constante y un posible sistema de coordenadas consiste en la posición de cada partícula, esto es, y_1 y y_2, pero este no es un buen sistema, ya que la constricción se expresa en él por

$$y_1 + y_2 = l. \tag{3.139}$$

Utilicemos como coordenadas generalizadas la posición con respecto a la polea de la primera partícula y y la longitud de la cuerda l. En estas coordenadas la constricción es $l = \text{cst.}$, por lo que son un buen sistema ya que la constricción se expresa como una coordenada constante. La transformación de estas coordenadas a las (y_1, y_2) es

$$y_1 = y, \tag{3.140}$$

$$y_2 = l - y. \tag{3.141}$$

Pensemos ahora en el sistema completo, esto es como si l pudiera variar, entonces la energía cinética del sistema sería

$$T = \frac{1}{2}m_1\dot{y}^2 + \frac{1}{2}m_2(\dot{y}^2 + \dot{l}^2 - 2\dot{y}\dot{l}). \tag{3.142}$$

Las fuerzas generalizadas, calculadas para el problema general, esto es incluyendo las fuerzas de constricción, que en este caso es sólo una, la tensión de la cuerda τ, se puede encontrar por el cálculo del trabajo necesario para hacer desplazamientos virtuales. Así para la coordenada y obtendremos

$$\delta W = (m_1 g - \tau)\delta y - (m_2 g - \tau)\delta y = (m_1 - m_2)g\delta y, \tag{3.143}$$

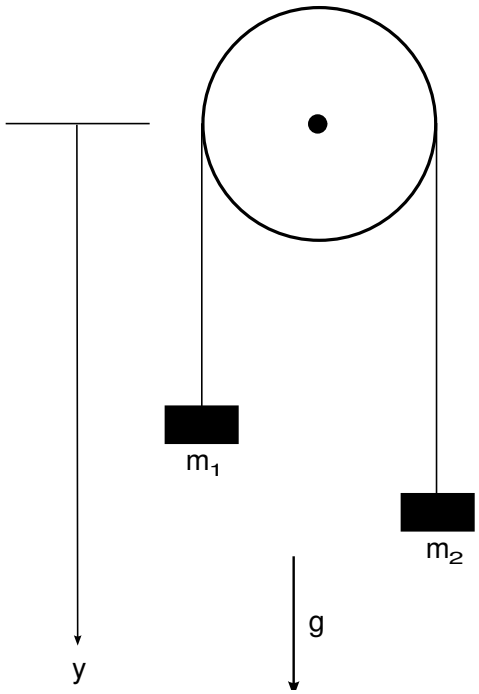


Figura 3.11: La máquina de Atwood.

donde, desde luego, hemos considerado que al variar y, dejando la coordenada l constante, se mueven ambas masas. De aquí encontramos que

$$Q_y = (m_1 - m_2)g. \tag{3.144}$$

Para la coordenada l tendremos

$$\delta W = (m_2 g - \tau)\delta l, \tag{3.145}$$

o sea que

$$Q_l = m_2 g - \tau. \tag{3.146}$$

Las ecuaciones de Lagrange (3.74) serán

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{y}}\right) - \frac{\partial T}{\partial y} = (m_1 + m_2)\ddot{y} = Q_y = (m_1 - m_2)g, \tag{3.147}$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{l}}\right) - \frac{\partial T}{\partial l} = -m_2\ddot{y} = Q_l = m_2 g - \tau. \tag{3.148}$$

De la primera ecuación obtenemos la ecuación de movimiento del sistema con constricción y de la segunda, al substituir la solución de la primera encontramos la fuerza de constricción τ. Los resultados son

$$y = y_0 + v_0 t + \frac{1}{2}\frac{m_1 - m_2}{m_1 + m_2} g t^2, \tag{3.149}$$

y

$$\tau = \frac{2m_1 m_2}{m_1 + m_2} g. \tag{3.150}$$

Consideremos otro ejemplo.

Un disco plano de radio r y masa m rueda, a partir del reposo, sobre un plano inclinado que hace un ángulo α con la horizontal. Deseamos conocer el mayor ángulo α para el cual, si el coeficiente de fricción estática entre el disco y el plano tiene un valor μ, el disco rueda sin resbalar, ver figura 3.12.

Cuando un disco plano en el plano está totalmente libre tiene tres grados de libertad, la posición sobre el plano de su centro y el ángulo de rotación. Como se ilustra en la figura utilizamos un sistema cartesiano (x, y) para localizar el centro, ponemos la coordenada x a lo largo del plano inclinado y el origen en un punto arbitrario de este, para especificar la rotación del disco utilizamos el ángulo θ que hace una marca arbitraria en el disco con la vertical.

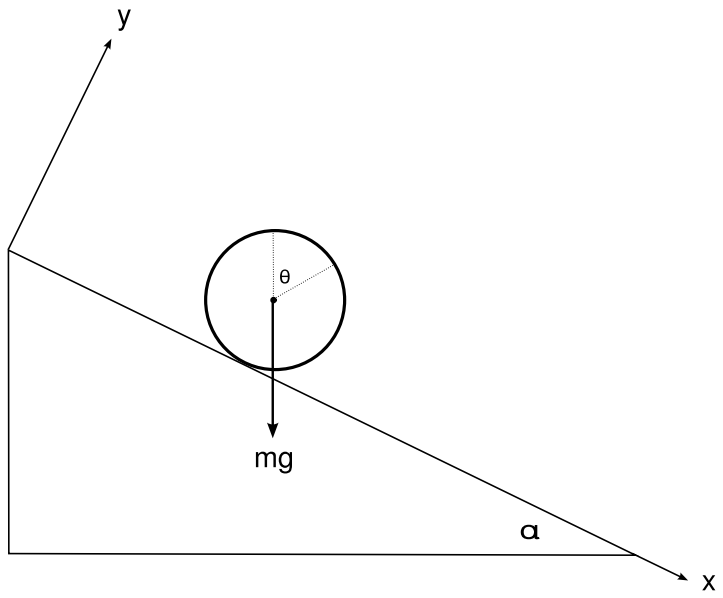


Figura 3.12: Un disco rueda sin resbalar sobre un plano inclinado.

El momento de inercia del disco con respecto a su centro es $mr^2/2$ por lo que la energía cinética del sistema será

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{4}mr^2\dot{\theta}^2. \tag{3.151}$$

Las constricciones son dos: que la distancia del centro del disco al plano sea su radio y que el disco ruede sin resbalar, en las coordenadas que hemos establecido se expresan por medio de

$$\begin{aligned} y &= r, \\ dx - rd\theta &= 0. \end{aligned} \tag{3.152}$$

Esta última es equivalente a

$$x - r\theta = c, \tag{3.153}$$

donde c es constante. Vemos entonces que las constricciones son holonómicas, pero las coordenadas no son adecuadas, ya que la segunda condición no está expresada por la constancia de una de las coordenadas. Para resolver este problema introducimos la coordenada $w = x - r\theta$ en substitución de θ, con lo

cual la constricción se expresara como $w = c$. En estas nuevas coordenadas la energía cinética será

$$T = \frac{3}{4}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + \frac{1}{4}m\dot{w}^2 - \frac{1}{2}m\dot{x}\dot{w}. \tag{3.154}$$

Calculemos ahora las fuerzas generalizadas incluyendo las fuerzas de constricción, que son la fuerza normal que ejerce el plano sobre el disco F_n y la fuerza de fricción F_f . Al hacer un desplazamiento virtual en x dejando las otras coordenadas fijas se realiza trabajo únicamente a favor de la componente x de la gravedad $mg\text{sen}\alpha$; podría pensarse que también se hace trabajo en contra de la fuerza de fricción, pero al tomar en cuenta que la coordenada w debe permanecer constante, al desplazar x por δx θ se desplaza por $\delta x/r$ y por tanto no se desplaza el punto de contacto y no se realiza trabajo contra la fuerza de fricción. Así

$$\delta W = \delta x mg\text{sen}\alpha, \tag{3.155}$$

por tanto

$$Q_x = mg\text{sen}\alpha. \tag{3.156}$$

Al desplazar virtualmente la coordenada y se hace trabajo a favor de la fuerza normal y en contra de la componente y de la gravedad y por tanto

$$Q_y = F_n - mg\cos\alpha. \tag{3.157}$$

Finalmente al desplazar virtualmente la coordenada w, manteniendo la x y la y constantes, solo se hace trabajo en contra de la fuerza de fricción, y

$$Q_w = -F_f. \tag{3.158}$$

Las ecuaciones de Lagrange serán

$$\begin{aligned} \frac{3}{2}m\ddot{x} - \frac{1}{2}m\ddot{w} &= mg\text{sen}\alpha, \\ m\ddot{y} &= F_n - mg\cos\alpha, \\ \frac{1}{2}m\ddot{w} - \frac{1}{2}m\ddot{x} &= -F_f. \end{aligned} \tag{3.159}$$

Queremos calcular las fuerzas F_n y F_f para mantener las constricciones, esto implica que y y w deben ser constantes y por tanto que $\ddot{y}$ y $\ddot{w}$ deben ser cero. Al imponer estas condiciones la primera ecuación, la correspondiente al único grado de libertad, nos da la ecuación de movimiento considerando las constricciones y, de las otras dos ecuaciones, al considerar la solución

$$x(t) = \frac{1}{3}gt^2\text{sen}\alpha + v_{0x}t + x_0 \tag{3.160}$$

de la primera,[8] obtenemos las fuerzas de constricción

$$\begin{aligned} F_n &= mg\cos\alpha, \\ F_f &= \frac{1}{3}mg\text{sen}\alpha. \end{aligned} \tag{3.161}$$

Estas serían las fuerzas necesarias para que el movimiento siguiera las constricciones. Ahora la fuerza de fricción no puede ser mayor que el coeficiente de fricción estática multiplicado por la fuerza normal, por tanto

$$\frac{1}{3}mg\text{sen}\alpha \leq \mu mg\cos\alpha, \tag{3.162}$$

de lo contrario mantener las constricciones exigiría una fuerza de fricción mayor que la máxima posible y el disco comenzaría a resbalar. Esto ocurrirá cuando el ángulo α sea tal que se cumpla la igualdad en la ecuación (3.162) , esto es, el disco resbala cuando

$$\alpha > \arctan(3\mu). \tag{3.163}$$

3.7. Sistemas no holonómicos.

Hemos visto que la formulación lagrangiana nos permite reducir la dimensionalidad con la que tratamos un sistema mecánico holonómico a la del número de sus grados de libertad y con esto simplificar su solución. Cuando el sistema es no holonómico esto no es posible. Sin embargo es posible atacar el estudio de los sistemas no holonómicos por la vía del llamado método de los multiplicadores indeterminados de Lagrange, veamos esto.

Comencemos con el problema de una partícula de masa m sujeta a una constricción. En general una constricción será una condición sobre los posibles desplazamientos de la partícula que se expresa, en coordenadas cartesianas (x, y, z), por

$$\alpha(x,y,z,t)dx + \beta(x,y,z,t)dy + \gamma(x,y,z,t)dz + \tau(x,y,z,t)dt = 0, \tag{3.164}$$

donde α, β, γ y τ son funciones del espacio y el tiempo que supondremos continuas y derivables tantas veces como sea necesario. Si esta forma diferencial

[8] En este caso no es necesario resolver la ecuación de movimiento en los grados de libertad, ya que las ecuaciones correspondientes a las constricciones únicamente dependen de $\ddot{x}$ y está puede ser calculada directamente. Esto se debe a la simplicidad del sistema, en general las ecuaciones correspondientes a las constricciones dependen de las velocidades y las coordenadas generalizadas correspondientes a los grados de libertad, en esos casos es necesario resolver las ecuaciones de movimiento para hacer las substituciones.

es integrable, esto es si existe un factor integrador para (3.164), entonces la constricción se puede expresar como

$$f(x, y, z, t) = cst \tag{3.165}$$

y será holonómica, en caso contrario será no holonómica. Para tratar el caso general, tanto holonómico como no holonómico, lo haremos a partir de la expresión diferencial de la constricción.

Un desplazamiento físico de la partícula será aquel que resulte del movimiento real, esto es aquel en que el desplazamiento (dx, dy, dz) sea resultado de haber transcurrido un intervalo de tiempo dt y, claro está, estas diferenciales deberán cumplir con (3.164).

Llamaremos a un desplazamiento virtual cuando, pensando en un instante fijo del tiempo, se desplaza a la partícula de manera arbitraria pero respetando las constricciones. Si denotamos a un desplazamiento virtual por $(\delta x, \delta y, \delta z)$ tendremos que debe cumplirse con

$$\alpha\delta x + \beta\delta y + \gamma\delta z = 0. \tag{3.166}$$

La fuerza total que actúa sobre la partícula puede dividirse en dos, las fuerzas aplicadas y las de constricción, estas últimas son las que obligarán a la partícula, a lo largo de su movimiento, a cumplir con las constricciones.

Así las ecuaciones de movimiento serán

$$\begin{aligned} m\ddot{x} &= F_x + C_x, \\ m\ddot{y} &= F_y + C_y, \\ m\ddot{z} &= F_z + C_z, \end{aligned} \tag{3.167}$$

donde las F son las componentes de la fuerza aplicada y las C las de la fuerza de constricción.

En general no es posible conocer a priori la fuerza de constricción pues depende del movimiento mismo que esté realizando la partícula, esto implica que las ecuaciones (3.167) no constituyen un sistema completo de ecuaciones pues desconocemos C_x, C_y y C_z. Por otro lado conocemos una propiedad básica de las fuerzas de constricción, el hecho de que no realizan trabajo alguno ante un desplazamiento virtual, esto es

$$C_x\delta x + C_y\delta y + C_z\delta z = 0 \tag{3.168}$$

cuando δx, δy y δz cumplen con (3.166). Esta propiedad nos permitirá resolver el problema de la incompletez de las ecuaciones (3.167).

Si se ha de cumplir (3.168) cuando se cumple (3.166) entonces tendremos que

$$\frac{C_x}{\alpha} = \frac{C_y}{\beta} = \frac{C_z}{\gamma}, \tag{3.169}$$

y llamando λ a esta razón tendremos

$$C_x = \lambda\alpha, \quad C_y = \lambda\beta \quad y \quad C_z = \lambda\gamma, \tag{3.170}$$

donde λ es un coeficiente por determinar o indeterminado.

Las ecuaciones de movimiento se pueden ahora escribir como

$$\begin{aligned} m\ddot{x} &= F_x + \lambda\alpha, \\ m\ddot{y} &= F_y + \lambda\beta, \\ m\ddot{z} &= F_z + \lambda\gamma, \end{aligned}$$

que son tres ecuaciones en cuatro variables x, y, z, y λ, pero tenemos una cuarta ecuación, pues las fuerzas de constricción son tales que obligan a la partícula a seguir la constricción por lo que debemos agregar la condición (3.164) y tendremos que

$$\alpha\dot{x} + \beta\dot{y} + \gamma\dot{z} + \tau = 0. \tag{3.171}$$

Ahora tenemos cuatro ecuaciones para determinar cuatro variables, por lo que podemos resolver las ecuaciones y, dadas unas condiciones iniciales, obtenemos $x(t)$, $y(t)$ y $z(t)$, esto es, el movimiento de la partícula; pero en adición obtendremos $\lambda(t)$ y de (3.170) las fuerzas de constricción en función del tiempo, hemos así resuelto el sistema completo, incluyendo el valor de las fuerzas de constricción.

Esto puede generalizarse al caso en que tengamos dos constricciones

$$\alpha_1(x,y,z,t)dx+\beta_1(x,y,z,t)dy+\gamma_1(x,y,z,t)dz+\tau_1(x,y,z,t)dt = 0, \tag{3.172}$$

$$\alpha_2(x,y,z,t)dx+\beta_2(x,y,z,t)dy+\gamma_2(x,y,z,t)dz+\tau_2(x,y,z,t)dt = 0. \tag{3.173}$$

En este caso se deberá cumplir que

$$C_x\delta x + C_y\delta y + C_z\delta z = 0 \tag{3.174}$$

cuando

$$\alpha_1\delta x + \beta_1\delta y + \gamma_1\delta z = 0, \tag{3.175}$$

$$\alpha_2\delta x + \beta_2\delta y + \gamma_2\delta z = 0. \tag{3.176}$$

Esto lo logramos introduciendo dos multiplicadores por determinar λ_1 y λ_2 y escribiendo las fuerzas de constricción como

$$\begin{aligned} C_x &= \lambda_1\alpha_1 + \lambda_2\alpha_2, \\ C_y &= \lambda_1\beta_1 + \lambda_2\beta_2, \\ C_z &= \lambda_1\gamma_1 + \lambda_2\gamma_2. \end{aligned} \tag{3.177}$$

Las ecuaciones de movimiento serán

$$\begin{aligned} m\ddot{x} &= F_x + \lambda_1\alpha_1 + \lambda_2\alpha_2, \\ m\ddot{y} &= F_y + \lambda_1\beta_1 + \lambda_2\beta_2, \\ m\ddot{z} &= F_z + \lambda_1\gamma_1 + \lambda_2\gamma_2, \end{aligned}$$

que son tres y contienen cinco variables x, y, z, λ_1 y λ_2; pero ahora tenemos dos ecuaciones adicionales, las dos constricciones

$$\begin{aligned} \alpha_1\dot{x} + \beta_1\dot{y} + \gamma_1\dot{z} + \tau_1 &= 0, \\ \alpha_2\dot{x} + \beta_2\dot{y} + \gamma_2\dot{z} + \tau_2 &= 0, \end{aligned}$$

por lo que en realidad tenemos cinco ecuaciones y cinco variables. Podemos, entonces, resolver el sistema y obtener, para unas condiciones iniciales, el movimiento y los dos multiplicadores, a partir de estos y de (3.177) podemos obtener las fuerzas de constricción.

Este procedimiento puede ser extendido a un sistema general de n partículas, cada una sujeta a una fuerza aplicada y a una fuerza de constricción que asegura que se cumplan una serie de constricciones. Tendremos entonces un espacio de configuración de $3n$ dimensiones, un sistema de coordenadas inducidas por un sistema cartesiano en el espacio físico y, en adición, digamos k constricciones. Las ecuaciones de movimiento serán

$$\begin{aligned} m_i\ddot{x}_i &= F_{ix} + C_{ix}, \\ m_i\ddot{y}_i &= F_{iy} + C_{iy}, \qquad\qquad i = 1, \ldots, n \\ m_i\ddot{z}_i &= F_{iz} + C_{iz}, \end{aligned}$$

donde las F son las componentes de las fuerzas aplicadas y las C las de constricción.

Las constricciones se expresarán por

$$\sum_{i=1}^{n} (\alpha_{li}dx_i + \beta_{li}dy_i + \gamma_{li}dz_i) + \tau_l dt = 0 \qquad\qquad l = 1, \ldots, k, \tag{3.178}$$

donde los coeficientes son funciones de las posiciones de las partículas y del tiempo.

El número de grados de libertad será, por definición, $3n-k$ y, ante un desplazamiento virtual del sistema se debe cumplir con

$$\sum_{i=1}^{n}(\alpha_{li}\delta x_i+\beta_{li}\delta y_i+\gamma_{li}\delta z_i)=0, \qquad l=1,\ldots,k, \tag{3.179}$$

la relación que se establece entre estas deltas debe implicar que

$$\sum_{i=1}^{n}(C_{ix}\delta x_i+C_{iy}\delta y_i+C_{iz}\delta z_i)=0. \tag{3.180}$$

Para hacer que se cumpla esta última ecuación cuando se cumplen las constricciones introducimos k multiplicadores λ_l por medio de

$$\begin{aligned} C_{ix} &= \sum_{l=1}^{k}\lambda_l\alpha_{li}, \\ C_{iy} &= \sum_{l=1}^{k}\lambda_l\beta_{li}, \\ C_{iz} &= \sum_{l=1}^{k}\lambda_l\gamma_{li} \end{aligned}$$

y las ecuaciones de movimiento serán

$$\begin{aligned} m_i\ddot{x}_i &= F_{ix}+\sum_{l=1}^{k}\lambda_l\alpha_{li}, \\ m_i\ddot{y}_i &= F_{iy}+\sum_{l=1}^{k}\lambda_l\beta_{li}, \\ m_i\ddot{z}_i &= F_{iz}+\sum_{l=1}^{k}\lambda_l\gamma_{li}, \end{aligned}$$

las que adicionadas de las k ecuaciones de constricción

$$\sum_{i=1}^{n}(\alpha_{li}\dot{x}_i+\beta_{li}\dot{y}_i+\gamma_{li}\dot{z}_i)+\tau_l=0, \tag{3.181}$$

establecen un sistema de $3n+k$ ecuaciones con $3n+k$ variables. Al resolver este sistema obtenemos el movimiento y los multiplicadores, los que, a su vez, nos proporcionan las fuerzas de constricción.

Para poner todo esto en términos lagrangianos podemos seguir, casi textualmente, los pasos que nos llevaron a la formulación lagrangiana al principio de este capítulo. Todo aquello que tiene que ver con la transformación de la energía cinética a coordenadas generalizadas, las definiciones de velocidades e impulsos generalizados y los cálculos que no involucran a las fuerzas, será idéntico a lo ya hecho. Con respecto a las fuerzas ahora tenemos las aplicadas y las de constricción, pero podemos podemos limitar la definición de fuerzas generalizadas a las fuerzas aplicadas, esto es

$$Q_j = \sum_{i=1}^{n} \left[F_{x_i} \frac{\partial x_i}{\partial q^j} + F_{y_i} \frac{\partial y_i}{\partial q^j} + F_{z_i} \frac{\partial z_i}{\partial q^j} \right] \qquad j = 1, \ldots, 3n \tag{3.182}$$

y adicionar unas fuerzas generalizadas de constricción

$$Q_j^c = \sum_{i=1}^{n} \left[C_{x_i} \frac{\partial x_i}{\partial q^j} + C_{y_i} \frac{\partial y_i}{\partial q^j} + C_{z_i} \frac{\partial z_i}{\partial q^j} \right], \tag{3.183}$$

llegaríamos a unas ecuaciones de Lagrange de la forma

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^j} \right) - \frac{\partial T}{\partial q^j} = Q_j + Q_j^c \tag{3.184}$$

o en caso de existir un potencial para las fuerzas aplicadas

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j} \right) - \frac{\partial L}{\partial q^j} = Q_j^c, \tag{3.185}$$

donde debemos recordar que las fuerzas generalizadas Q_j^c no son conocidas a priori por no serlo las de constricción.

Ahora haciendo la transformación a coordenadas generalizadas de las expresiones para las constricciones llegaremos a las expresiones

$$\sum_{j=1}^{3n} A_{lj} dq^j + T_l dt = 0, \qquad l = 1, \ldots, k \tag{3.186}$$

donde

$$A_{lj} = \sum_{i=1}^{n} \left(\alpha_{li} \frac{\partial x_i}{\partial q^j} + \beta_{li} \frac{\partial y_i}{\partial q^j} + \gamma_{li} \frac{\partial z_i}{\partial q^j} \right) \tag{3.187}$$

y

$$T_l = \sum_{i=1}^{n} \left(\alpha_{li} \frac{\partial x_i}{\partial t} + \beta_{li} \frac{\partial y_i}{\partial t} + \gamma_{li} \frac{\partial z_i}{\partial t} \right) + \tau_l. \tag{3.188}$$

Ante un desplazamiento virtual tendremos que

$$\delta q^j = \sum_{i=1}^{n} \left(\frac{\partial q^j}{\partial x_i} \delta x_i + \frac{\partial q^j}{\partial y_i} \delta y_i + \frac{\partial q^j}{\partial z_i} \delta z_i \right), \tag{3.189}$$

con las condiciones

$$\sum_{j=1}^{3n} A_{lj} \delta q^j = 0 \qquad l = 1, \ldots, k. \tag{3.190}$$

Es fácil hacer ver que el hecho de que ante un desplazamiento virtual las fuerzas de constricción no realizan trabajo, se expresa en coordenadas generalizadas por

$$\sum_{j=1}^{3n} Q_j^c \delta q^j = 0. \tag{3.191}$$

Al unir todas estas ecuaciones tendremos una forma lagrangiana de expresar las ecuaciones de movimiento, las de constricción y la condición sobre la nulidad del trabajo realizado por las fuerzas de constricción ante un desplazamiento virtual.

Podemos ver que la situación es totalmente análoga al caso en coordenadas cartesianas, por lo que podemos utilizar el método de los multiplicadores indeterminados para encontrar la solución del sistema. Definimos k multiplicadores λ_l y hacemos

$$Q_j^c = \sum_{l=1}^{k} \lambda_l A_{lj}. \tag{3.192}$$

Las ecuaciones de Lagrange serán

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}^j} \right) - \frac{\partial T}{\partial q^j} = Q_j + \sum_{l=1}^{k} \lambda_l A_{lj} \tag{3.193}$$

o, si existe un potencial para las fuerzas aplicadas

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^j} \right) - \frac{\partial L}{\partial q^j} = \sum_{l=1}^{k} \lambda_l A_{lj}, \tag{3.194}$$

a estas ecuaciones hay que adicionar las de constricción expresadas por

$$\sum_{j=1}^{3n} A_{lj}\dot{q}^j + T_l = 0, \tag{3.195}$$

con esto tenemos tantas ecuaciones como variables y al resolverlas obtenemos el movimiento expresado por medio de $q_j(t)$ y los multiplicadores de Lagrange; a partir de ellos podemos calcular las fuerzas generalizadas de constricción y de ellas las fuerzas de constricción.

Desde luego si una o varias de las constricciones son holonómicas, podemos ir a un sistema de coordenadas en las que las constricciones holonómicas se expresen por la constancia de las correspondientes coordenadas y reducir el número de coordenadas y de ecuaciones, simplificando el problema. Naturalmente de proceder así no podremos calcular todas las fuerzas de constricción.

Como un ejemplo del uso de este método para calcular fuerzas de constricción, resolveremos el caso del disco que rueda sin resbalar por un riel inclinado, ya resuelto en la sección anterior, ver figura 3.12.

En este caso cualquier sistema de coordenadas es adecuado, por lo que utilizaremos las coordenadas (x, y, θ). La energía cinética será

$$T = \frac{1}{2}m(\dot{x} + \dot{y}) + \frac{1}{4}mr^2\dot{\theta}^2, \tag{3.196}$$

y el potencial

$$V = -mgx\text{sen}\alpha + mgy\cos\alpha, \tag{3.197}$$

por lo que la lagrangiana es

$$L = \frac{1}{2}m(\dot{x} + \dot{y}) + \frac{1}{4}mr^2\dot{\theta}^2 + mgx\text{sen}\alpha - mgy\cos\alpha. \tag{3.198}$$

Las constricciones se expresan en forma diferencial por

$$dy = 0, \tag{3.199}$$

$$dx - rd\theta = 0. \tag{3.200}$$

Asociaremos a la primera un multiplicador de Lagrange λ_1 y a la segunda otro multiplicador λ_2.

Las ecuaciones (3.194) serán

$$m\ddot{x} - mg\sin\alpha = \lambda_2, \tag{3.201}$$

$$m\ddot{y} + mg\cos\alpha = \lambda_1, \tag{3.202}$$

$$\frac{1}{2}r^2 m\ddot{\theta} = -r\lambda_2. \tag{3.203}$$

Las ecuacines de constricción (3.195) quedan como

$$\dot{y} = 0 \tag{3.204}$$

y

$$\dot{x} - r\dot{\theta} = 0. \tag{3.205}$$

Tenemos así cinco ecuaciones en cinco incógnitas.

De la tercera ecuación encontramos que

$$y = r = \text{cst}. \tag{3.206}$$

y por tanto $\ddot{y} = 0$, esto implica usando la segunda ecuación que

$$\lambda_1 = mg\cos\alpha. \tag{3.207}$$

Derivando la última de las ecuaciones obtenemos que $\ddot{\theta} = \ddot{x}/r$ y al substituir esto en la tercera obtenemos $\lambda_2 = -m\ddot{x}/2$, por lo que la primera ecuación se reduce a

$$\frac{3}{2}\ddot{x} = g\text{sen}\alpha \tag{3.208}$$

y entonces

$$x = \frac{1}{3}gt^2\text{sen}\alpha + v_{0x}t + x_0, \tag{3.209}$$

$$\lambda_2 = -\frac{1}{3}mg\text{sen}\alpha \tag{3.210}$$

y

$$\theta = \frac{1}{r}\left[\frac{1}{3}gt^2\text{sen}\alpha + v_{0x}t + r\theta_0\right]. \tag{3.211}$$

Hemos así encontrado el movimiento y los dos coeficientes de Lagrange. De estos últimos podemos calcular las fuerzas de constricción usando (3.192)

$$C_x = \lambda_2 = -\frac{1}{3}mg\sin\alpha, \tag{3.212}$$

$$C_y = \lambda_1 = mg\cos\alpha \tag{3.213}$$

y

$$C_\theta = -r\lambda_2 = \frac{r}{3}mg\sin\alpha. \tag{3.214}$$

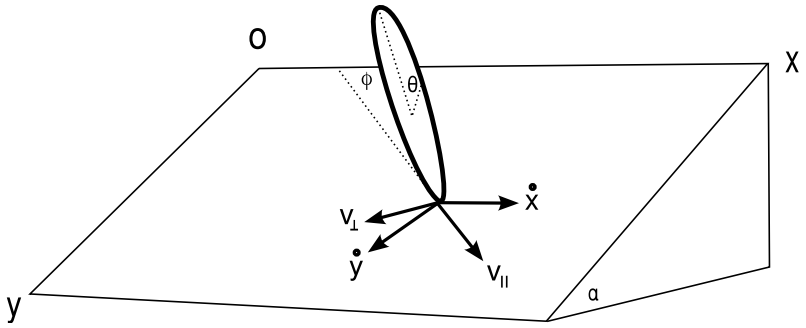


Figura 3.13: Un disco rueda sin resbalar sobre un plano inclinado.

Por simple inspección vemos que C_x es menos la fuerza de fricción, C_y en la fuerza normal y $-C_\theta$ es r veces la de fricción, con lo que hemos resuelto totalmente el sistema y reproducido los resultados obtenidos en el inciso anterior.

Veamos ahora un ejemplo de un sistema no holonómico: un disco homogéneo de masa m y radio r que rueda sin resbalar sobre un plano inclinado por un ángulo α mientras permanece perpendicular a dicho plano, como se ilustra en la figura 3.13.

La energía cinética está dada por

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{4}mr^2\dot{\theta}^2 + \frac{1}{8}mr^2\dot{\phi}^2, \tag{3.215}$$

el potencial por

$$V = -mgy\text{sen}\alpha \tag{3.216}$$

y la lagrangiana será

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{4}mr^2\dot{\theta}^2 + \frac{1}{8}mr^2\dot{\phi}^2 + mgy\text{sen}\alpha. \tag{3.217}$$

En forma diferencial las constricciones están dadas por la expresión (3.71)

$$\begin{aligned} dx\cos\phi + dy\text{sen}\phi - rd\theta &= 0, \\ -dx\text{sen}\phi + dy\cos\phi &= 0. \end{aligned}$$

Asignanado el multiplicador λ_1 a la primera constricción y λ_2 a la segunda tendremos las ecuaciones de movimiento

$$m\ddot{x} = \lambda_1\cos\phi - \lambda_2\text{sen}\phi, \tag{3.218}$$

$$m\ddot{y} - mg\text{sen}\alpha = \lambda_1 \text{sen}\phi + \lambda_2 \cos\phi, \tag{3.219}$$

$$\frac{1}{2}mr^2\ddot{\theta} = -r\lambda_1, \tag{3.220}$$

$$\frac{1}{4}mr^2\ddot{\phi} = 0, \tag{3.221}$$

las que aunadas a las ecuaciones de constricción

$$\dot{x}\cos\phi + \dot{y}\text{sen}\phi - r\dot{\theta} = 0 \tag{3.222}$$

y

$$-\dot{x}\text{sen}\phi + \dot{y}\cos\phi = 0, \tag{3.223}$$

constituyen un sistema completo de seis ecuaciones en seis incógnitas.

De las primeras dos ecuaciones podemos despejar λ_1 y λ_2, obtenemos

$$\lambda_1 = m\left[\ddot{x}\cos\phi + \ddot{y}\text{sen}\phi - g\text{sen}\alpha\text{sen}\phi\right] \tag{3.224}$$

y

$$\lambda_2 = m\left[-\ddot{x}\text{sen}\phi + \ddot{y}\cos\phi - g\text{sen}\alpha\cos\phi\right]. \tag{3.225}$$

Ahora derivando las ecuaciones de constricción obtenemos

$$\ddot{x}\cos\phi - \dot{x}\dot{\phi}\text{sen}\phi + \ddot{y}\text{sen}\phi + \dot{y}\dot{\phi}\cos\phi - r\ddot{\theta} = 0 \tag{3.226}$$

y

$$-\ddot{x}\text{sen}\phi - \dot{x}\dot{\phi}\cos\phi + \ddot{y}\cos\phi - \dot{y}\dot{\phi}\text{sen}\phi = 0, \tag{3.227}$$

por lo que tendremos

$$\lambda_1 = m\left[\dot{x}\dot{\phi}\text{sen}\phi - \dot{y}\dot{\phi}\cos\phi - g\text{sen}\alpha\text{sen}\phi\right] \tag{3.228}$$

y

$$\lambda_2 = m\left[\dot{x}\dot{\phi}\cos\phi + \dot{y}\dot{\phi}\text{sen}\phi - g\text{sen}\alpha\cos\phi\right], \tag{3.229}$$

introduciendo estos valores en las ecuaciones de movimiento tendremos

$$\ddot{x} = -\dot{y}\dot{\phi} + r\ddot{\theta}\cos\phi, \tag{3.230}$$

$$\ddot{y} = \dot{x}\dot{\phi} + r\ddot{\theta}\text{sen}\phi, \tag{3.231}$$

$$\ddot{\theta} = -\frac{2}{3r}\left[\dot{x}\dot{\phi}\text{sen}\phi - \dot{y}\dot{\phi}\cos\phi - g\text{sen}\alpha\text{sen}\phi\right] \tag{3.232}$$

y

$$\ddot{\phi} = 0, \tag{3.233}$$

que ya constituyen un sistema completo.

Como era de esperar $\dot{\phi}$ es una constante y, si inicialmente vale cero, las ecuaciones se reducen y el movimiento del centro de masas es sobre una recta.

3.8. El principio de Hamilton.

Pasemos ahora a ver la mecánica clásica desde otra perspectiva, la de los principios variacionales. Para hacer esto es necesaria una pequeña introducción al cálculo de variaciones.

El cálculo diferencial ordinario se construye para funciones o mapeos de R^n en R^m con la norma euclidiana, esto es, el espacio R^n se entiende como espacio vectorial y la norma o "tamaño" de un vector $(x_1, x_2, \cdots, x_n)$ se define como la raíz de $(x_1^2 + x_2^2 + \cdots + x_n^2)$. Esto se hace así debido a que, por una parte, la norma nos permite dar una estructura topológica y definir la convergencia de las sucesiones de puntos para llegar por esta vía al concepto de límite, y por otra parte, el estructurar R^n como espacio vectorial nos permite definir la derivada de un mapeo valuada en un punto, como el mapeo lineal que mejor aproxima al mapeo original en ese punto, y el concepto de linealidad no se puede dar mas que en espacios vectoriales.

En general el cálculo diferencial puede extenderse a los espacios de Banach, esto es espacios vectoriales (dimensión finita o infinita), con norma y completos ante la métrica inducida por la norma[9], ejemplos importantes son el propio R^n con la norma euclidiana y los espacios de Hilbert de la mecánica cuántica.

Aquí estamos interesados en las funcionales. Una funcional es un mapeo de un conjunto de curvas en R^n en R, esto es, la asociación de un número real a cada curva, las que, recordamos, son funciones de R en R^n. Un ejemplo es la longitud de las curvas dadas por $(x_i(s))$, esto es mapeos de R en R^n parametrizados por s. Tomemos dos valores de s digamos s_0 y s_1, definimos la funcional como la longitud de la curva entre los puntos $(x_i(s_0))$ y $(x_i(s_1))$. La expresión para esta funcional es

$$\int_{s_0}^{s_1} \sqrt{\left(\frac{dx_1}{ds}\right)^2 + \left(\frac{dx_2}{ds}\right)^2 + \cdots + \left(\frac{dx_n}{ds}\right)^2}\, ds. \tag{3.234}$$

A los conjuntos de curvas en R^n es sencillo dotarlos de la estructura de un espacio vectorial normado y por tanto extender los conceptos del cálculo a las funcionales; en particular el de derivada o diferencial valuada en un punto, veamos esto.

Cuando tenemos un mapeo de R^n en R^m, digamos $f : R^n \longrightarrow R^m$, decimos que f es diferenciable en el punto $x \in R^n$, si ocurre que en una

[9] Un espacio métrico es completo si toda sucesión de Cauchy es convergente a un punto del mismo espacio. Podemos decir que casi todos los espacios métricos de interés en la física son completos.

vecindad de x, f puede aproximarse por un mapeo lineal de R^n en R^m; esto es si ocurre que

$$f(x+h) - f(x) = Df_x[h] + O(|h|^2), \tag{3.235}$$

donde $Df_x : R^n \longrightarrow R^m$ denota un mapeo lineal y con el segundo sumando del miembro derecho de la ecuación denotamos términos que tienden a cero con h como, o más rápido que, el cuadrado de la norma de h. Al mapeo lineal Df_x se le llama la diferencial del mapeo f en el punto x; en coordenadas el mapeo lineal Df_x queda representado por la matriz jacobiana de f valuada en x.

En el caso de las funcionales podemos hacer algo similar. Dada una curva $c : R \longrightarrow R^n$ que pertenece a un espacio de Banach de curvas $\mathcal{C}$ y una funcional, esto es un mapeo $\mathcal{F} : \mathcal{C} \longrightarrow R$ que a cada curva le asocia un número real, diremos que la funcional es diferenciable en la curva c si para curvas cercanas a c, o sea en una vecindad de c, la funcional se puede aproximar por una funcional lineal. Esto es si dada una curva h ocurre que

$$\mathcal{F}(c+h) - \mathcal{F}(c) = D\mathcal{F}_c[h] + O(|h|^2), \tag{3.236}$$

en esta ocasión c y h son curvas y $D\mathcal{F}_c$ es una funcional lineal, pero el concepto es totalmente similar al caso anterior. A la funcional lineal $D\mathcal{F}_c$ se le llama la diferencial de la funcional $\mathcal{F}$ valuada en la curva c. Si la funcional es diferenciable en todas las curvas se dice simplemente que es diferenciable. Todas las propiedades de los mapeos diferenciables y los teoremas del cálculo diferencial común son aplicables al caso de las funcionales, por ejemplo la unicidad de la diferencial, esto es si una funcional es diferenciable en una curva, la diferencial valuada en esa curva es un mapeo lineal único.

Pasemos ahora a estudiar un tipo especial de funcionales. Pensemos en las curvas en R^n esto es mapeos $c : R \longrightarrow R^n$, llamemos $t \in R$ al parámetro y usemos un sistema de coordenadas cartesianas en R^n, entonces la curva c quedará descrita por $(c_1(t), c_2(t), \cdots, c_n(t))$ y la "velocidad" con que se recorre esta curva o, mejor dicho, su tangente será $(\dot{c}_1(t), \dot{c}_2(t), \cdots, \dot{c}_n(t))$. Tomemos ahora una función $F : R^n \times R^n \times R \longrightarrow R$, donde entenderemos por el primer R^n la curva c, por el segundo R^n la tangente o "velocidad" y por el último R el parámetro t; entonces para cada valor de t

$$F(c_1(t), c_2(t), \cdots, c_n(t); \dot{c}_1(t), \dot{c}_2(t), \cdots, \dot{c}_n(t); t) \tag{3.237}$$

es un número real y, para dos valores fijos de t, digamos t_0 y t_1, construimos la funcional siguiente

$$\mathcal{F}(c) = \int_{t_0}^{t_1} F(c(t), \dot{c}(t), t)dt. \tag{3.238}$$

Notemos que la longitud de la curva vista como funcional en la ecuación (3.234) es un caso particular de este tipo de funcionales.

Calculemos la diferencial de una funcional de este tipo, claro en el proceso se verá que se trata de una funcional diferenciable, de lo contrario no tendría sentido dicho cálculo. Utilicemos la definición dada por (3.236)

$$\begin{aligned}\mathcal{F}(c+h)-\mathcal{F}(c) &= \int_{t_0}^{t_1} F(c(t)+h(t),\dot{c}(t)+\dot{h}(t),t)dt \\ &\quad -\int_{t_0}^{t_1} F(c(t),\dot{c}(t),t)dt \\ &= \int_{t_0}^{t_1}\left[F\left(c(t)+h(t),\dot{c}(t)+\dot{h}(t),t\right)-F\left(c(t),\dot{c}(t),t\right)\right]dt \end{aligned} \tag{3.239}$$

Pero F es una función diferenciable del cálculo común de R^{n+n+1} en R y por tanto podemos escribir

$$\mathcal{F}(c+h)-\mathcal{F}(c)=\int_{t_0}^{t_1}\left[DF_c[h,\dot{h},t]+O(|h|^2)\right]dt \tag{3.240}$$

$$=\int_{t_0}^{t_1}\left[\sum_{i=1}^{n}\left(\frac{\partial F}{\partial c_i}h_i+\frac{\partial F}{\partial \dot{c}_i}\dot{h}_i\right)\Big|_c+O(|h|^2)\right]dt \tag{3.241}$$

$$=\sum_{i=1}^{n}\left[\int_{t_0}^{t_1}\frac{\partial F}{\partial c_i}\Big|_c h_i dt+\int_{t_0}^{t_1}\frac{\partial F}{\partial \dot{c}_i}\Big|_c \dot{h}_i dt\right]+O(|h|^2) \tag{3.242}$$

$$=\sum_{i=1}^{n}\left[\int_{t_0}^{t_1}\frac{\partial F}{\partial c_i}\Big|_c h_i dt-\int_{t_0}^{t_1}\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{c}_i}\right)\Big|_c h_i dt\right]+\frac{\partial F}{\partial \dot{c}_i}\Big|_c h_i\Big|_{t_0}^{t_1}+O(|h|^2) \tag{3.243}$$

$$=\sum_{i=1}^{n}\int_{t_0}^{t_1}\left[\frac{\partial F}{\partial c_i}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{c}_i}\right)\right]_c h_i dt+\sum_{i=1}^{n}\left[\frac{\partial F}{\partial \dot{c}_i}\Big|_c h_i\right]_{t_0}^{t_1}+O(|h|^2). \tag{3.244}$$

Donde hemos hecho uso de la integración por partes para pasar de la tercera a la cuarta expresión.

La última ecuación nos muestra que en efecto la funcional es diferenciable y que la diferencial está dada por

$$D\mathcal{F}_c[h]=\sum_{i=1}^{n}\int_{t_0}^{t_1}\left[\frac{\partial F}{\partial c_i}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{c}_i}\right)\right]_c h_i dt+\sum_{i=1}^{n}\left[\frac{\partial F}{\partial \dot{c}_i}\Big|_c h_i\right]_{t_0}^{t_1}. \tag{3.245}$$

En el cálculo de las funciones de R^n en R decimos que una función tiene un extremo en un cierto punto, si su diferencial en ese punto es nula; esto es, si aplicada como función lineal a cualquier punto da cero. En el caso de las funcionales podemos decir que una funcional tiene un extremo en una cierta curva si su diferencial valuada en esa curva es tal que al aplicarla a cualquier curva da cero. Naturalmente las extremales en ambos casos pueden representar máximos, mínimos o inflexiones en torno al punto o curva en que se da el valor extremo.

En muchas ocasiones a la h que hemos estado utilizando se le llama una variación, con esto se da entender que se tiene una curva c y se la varía por h. Así mismo al cambio a primer orden que ocurre en la funcional al hacer una variación de la curva y que queda expresado en la ecuación (3.244) se le llama variación de la funcional. En general se denota a la variación de la curva por δc o en coordenadas por δc_i; de la misma forma a la variación de la funcional se le denota por $\delta\mathcal{F}$. En este contexto al cálculo en el espacio de las funcionales se le llama cálculo de variaciones. Nótese que la variación de la funcional es lo que hemos llamado la diferencial de la funcional aplicada a la variación de la curva.

Tomemos como ejemplo la longitud de una curva en el espacio de dos dimensiones entre los valores t_0 y t_1 del parámetro que la define. Limitemos las posibles curvas a aquellas que para $t = t_0$ pasan por un punto fijo (x_0, y_0) y para el valor $t = t_1$ por otro punto (x_1, y_1); el cálculo de la extremal de la funcional (3.234) nos proporcionará la curva que hace extremal la longitud entre esos dos puntos. Como la funcional (3.234) es de la forma (3.238) la fórmula (3.245) nos da su diferencial y al exigir que ésta sea cero para toda h_i obtendríamos las ecuaciones diferenciales para la extremal, que de corresponder a un mínimo sería la geodésica que pasa por los dos puntos.

Ahora bien como en los valores t_0 y t_1 del parámetro todas las curvas deben pasar por los mismos dos puntos, en esos valores $h(t)$ debe ser cero, por lo que el segundo término de la ecuación (3.245) es cero y como el resto ha de ser cero independientemente de $h(t)$, entonces los términos dentro del paréntesis cuadrado deben ser cero, esto es

$$\frac{\partial F}{\partial c_i} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{c}_i}\right) = 0, \qquad i = 1, 2. \tag{3.246}$$

Estas dos ecuaciones al considerar que

$$F = \sqrt{\dot{x}^2 + \dot{y}^2} \tag{3.247}$$

nos dan

$$\begin{aligned} \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) &= 0, \\ \frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) &= 0, \end{aligned} \tag{3.248}$$

de aquí encontramos que

$$\frac{dx}{dy} = \text{cst.} \tag{3.249}$$

y por tanto que la curva que pasa por dos puntos y cuya longitud es extremal, en este caso un mínimo, es la recta que los une. Claro está, si en vez de utilizar una métrica euclidiana hubiésemos utilizado otra, entonces la expresión (3.247) sería otra, la correspondiente a esa métrica, y las ecuaciones diferenciales nos llevarían a las geodésicas en esa métrica, pero las ecuaciones (3.246) serían las mismas. Además estos cálculos no tienen por que limitarse a dos dimensiones, únicamente tendríamos que considerar tantas ecuaciones de la forma (3.246) como dimensiones y obtendríamos la geodésica correspondiente.

En general, dada una funcional de la forma (3.238), y dentro del espacio de las curvas que pasan por un punto específico para $t = t_0$ y por otro para $t = t_1$, las curvas para las que la funcional tiene una extremal están dadas por las ecuaciones

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{c}_i}\right) - \frac{\partial F}{\partial c_i} = 0 \qquad i = 1, \cdots, n. \tag{3.250}$$

Ahora estas ecuaciones serían las ecuaciones de Lagrange para un sistema mecánico si la función F fuese su función lagrangiana, cuestión totalmente posible dado que las funciones lagrangianas siempre tienen la forma (3.237). Esto implica que dado un sistema mecánico lagrangiano siempre podemos decir que sus trayectorias son aquellas curvas sobre las que la funcional

$$\int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)dt, \tag{3.251}$$

es extremal. A esta funcional se le llama la funcional de acción del sistema mecánico y se suele denotar con la letra $\mathcal{S}$.

Lo que hemos encontrado es que, dado un sistema mecánico lagrangiano con lagrangiana $L(q, \dot{q}, t)$, dos puntos del espacio o variedad de configuración q_0 y q_1 y dos tiempos t_0 y t_1, la trayectoria física que el sistema mecánico

seguirá bajo la condición de que al tiempo t_0 pase por el punto q_0 y al t_1 por q_1, es aquella curva que, pasando por esos dos puntos en esos tiempos, hace que la funcional de acción sea extrema. Esto es así porque ambas, la trayectoria física del sistema y la curva extremal, vienen dadas por las mismas ecuaciones: las ecuaciones de Lagrange.[10]

Puede pensarse que el especificar que la curva pase por dos puntos fijos de la variedad de configuración a dos tiempos específicos es una limitación, pero no lo es. Como hemos visto las ecuaciones diferenciales de la mecánica en el espacio de configuración son siempre de segundo orden; esto hace que para delimitar la trayectoria de un sistema tengamos que especificar condiciones iniciales en un número que es el doble del número de grados de libertad. Es costumbre hacerlo dando los valores de la posición y de la velocidad a un tiempo inicial arbitrario, esto es especificando n coordenadas generalizadas y n velocidades generalizadas al tiempo t_0; pero podemos hacer lo mismo especificando una posición, digamos inicial, al tiempo t_0 y una posición, digamos final, al tiempo t_1, esto implica especificar $2n$ coordenadas generalizadas, n al tiempo t_0 y otras n al tiempo t_1, a cambio de no especificar ninguna velocidad generalizada. Naturalmente ambas formas de especificar las "condiciones iniciales" para encontrar una solución de las ecuaciones de movimiento son equivalentes y podemos, en principio, deducir unas a partir de las otras[11].

Podemos, pues, formular la mecánica diciendo que: Dado un sistema mecánico lagrangiano, su movimiento será aquel en que la funcional de acción, definida por la expresión (3.251), adquiera valores extremales. Esto se conoce como el principio de Hamilton o principio de acción extrema.

En general a la formulación de alguna teoría física en términos de extremales de funcionales, se le llama formulación variacional o en base a principios variacionales, el principio de Hamilton es un principio variacional para la mecánica.

---

[10] En el cálculo variacional o de funcionales a las ecuaciones (3.250) se les llama ecuaciones de Euler. En ocasiones se utiliza, tanto en la mecánica como en el cálculo de variaciones, la nomenclatura "ecuaciones de Euler-Lagrange", aquí simplemente las llamaremos ecuaciones de Lagrange.

[11] Para algunos sistemas mecánicos existen varias trayectorias que pasan por dos puntos determinados a tiempos determinados. Esto obedece al hecho de que los extremos, tanto para funciones como para funcionales, son sólo locales y no globales y, así como una función puede tener varios máximos y mínimos, una funcional puede tener varias curvas extremales, esto aún en el caso en que éstas deban pasar por dos puntos específicos. Naturalmente todas las extremales que tenga serán trayectorias físicas válidas ya que satisfacen a las ecuaciones de Lagrange; ocurre que tendremos varias condiciones iniciales, en el sentido de especificar un punto y unas velocidades generalizadas a un tiempo inicial, para una sola condición expresada como el paso de la trayectoria por dos puntos determinados a tiempos determinados.

Al formular la mecánica en términos de las ecuaciones de Lagrange introducimos un espacio abstracto, el espacio o variedad de configuración, así logramos representar su situación geométrica por un punto de esa variedad y pudimos contemplar con mayor amplitud el movimiento del sistema. Esto nos abrió nuevas perspectivas que nos permitieron profundizar conceptualmente en la mecánica, concentrar la atención en cuestiones que resultan importantes, como dejar a un lado las fuerzas y dimensiones de constricción u obtener ventajas en la solución de problemas concretos. Lo mismo ocurre al formular la mecánica a partir del principio de Hamilton, nos proporciona un nuevo punto de vista, nos permite profundizar y ligar la mecánica con otras ramas de la física, por ejemplo con la óptica por medio del principio de Fermat; pero es sólo otra forma de formular la mecánica que, como hemos visto, es totalmente equivalente a las ecuaciones de Lagrange. En muchas ocasiones el principio de Hamilton, en especial cuando se le da la forma de un principio de mínima acción, se ha utilizado para introducir especulaciones casi místicas sobre el comportamiento de la naturaleza. Es importante no caer en estas especulaciones que nada tienen que ver con la física; el principio de Hamilton es tan valido como la leyes de Newton, su justificación es la misma, el experimento, el desarrollo teórico y la concordancia con innumerables experiencias sustentadas en la realidad física.

3.9. Simetrías e integrales de movimiento.

En varias ocasiones hemos mencionado la relación que existe entre las simetrías de un sistema mecánico y sus integrales de movimiento. Esta relación queda establecida en un resultado general conocido como teorema de Noether, veamos esto con cuidado.

En primer término debemos definir con precisión que entendemos por una simetría de un sistema mecánico. Decimos que un objeto es simétrico si queda invariante al modificarlo de alguna forma específica, por ejemplo una esfera queda invariante ante cualquier rotación por un eje que pase por su centro, por ello decimos que, cualquier objeto invariante ante rotaciones por ejes que pasan por un punto, tiene simetría esférica. En el caso de los sistemas mecánicos las simetrías de que hablamos son aquellas que se presentan cuando el sistema queda invariante ante transformaciones en el espacio de configuración. Comencemos entonces por estudiar las transformaciones del espacio o variedad de configuración, para poder después definir la invariancia del sistema ante dichas transformaciones.

Un grupo de un parámetro de transformaciones del espacio o variedad de

configuración M en si misma es una colección, parametrizada continuamente por un parámetro real α, de mapeos diferenciables biunívocos con inversos diferenciables $F_\alpha : M \to M$ de esa variedad de configuración en si misma[12], que cumplen con las siguientes propiedades:

1. Al valor cero del parámetro le corresponde la transformación idéntica.
2. Dados dos mapeos de la colección F_α y F_β, el mapeo compuesto $F_\alpha \circ F_\beta$ es el mapeo $F_{\alpha+\beta}$, que naturalmente pertenece a la colección.
3. Todo mapeo F_α de la colección tiene un inverso en la colección que está dado por $F_{-\alpha}$.
4. La operación de composición de mapeos de la colección es asociativa[13].

Esto es la colección forma un grupo ante la composición de mapeos y la estructura del grupo es la del grupo de los reales ante la suma, esto es, el grupo de mapeos es isomorfo al de los reales.

Una forma de tratar con estas transformaciones o deformaciones del espacio consiste en contemplarlas como transformaciones de coordenadas, esto significa, por ejemplo, estudiar las rotaciones del espacio por un ángulo α en torno a un cierto eje, poniendo coordenadas en el espacio y haciendo transformaciones de coordenadas consistentes en rotar el sistema de coordenadas en torno al mismo eje y por un ángulo α. Notamos que una vez puesto el sistema de coordenadas las fórmulas que en él describen el mapeo M_α, esto es, las fórmulas que, dadas las coordenadas de un punto, nos proporcionan las del punto transformado, serán las mismas que representan la transformación de coordenadas, esto es, las que dadas las coordenadas de un punto nos proporcionan las coordenadas de ese mismo punto en el sistema de coordenadas transformado. Vemos pues que, una vez puesto un sistema de coordenadas, ambas formas de contemplar el asunto son equivalentes. Naturalmente al tratar con coordenadas debemos considerar que sólo tratamos con cuestiones locales, si queremos abarcar la totalidad de la variedad debemos recurrir a la teoría de las variedades diferenciales.

Notemos que al decir que tratamos con transformaciones diferenciables de la variedad de configuración en si misma estamos implicando que la variedad misma queda invariante, esto es, si es una esfera seguirá siendo una esfera después de haber realizado la transformación, esto es importante cuando

[12] Esto es, una colección parametrizada por α de difeomorfismos de la variedad en si misma.

[13] La operación de composición de mapeos es siempre asociativa.

utilizamos coordenadas pues al ser, como hemos dicho, coordenadas locales debemos tener cuidado que al realizar transformaciones de coordenadas éstas lo sean de verdad, esto es que sean diferenciables e invertibles y que no presenten ninguna singularidad, por ejemplo, el paso de coordenadas cartesianas a polares queda excluido como un cambio de coordenadas si queremos verlo como equivalente a un difeomorfismo del plano en si mismo.

Pasemos ahora a definir la invariancia de un sistema mecánico ante un grupo de un parámetro de transformaciones de su espacio de configuración.

Un sistema mecánico consiste de dos partes, un espacio o variedad de configuración y una lagrangiana. La primera parte refleja la geometría del sistema, la segunda la dinámica. Diremos que un sistema mecánico queda invariante ante una transformación o ante un grupo de transformaciones de un parámetro de su variedad de configuración, si su lagrangiana es invariante ante dicha transformación o ante todas las transformaciones del grupo; la invariancia de la variedad de configuración queda implícita por tratarse de difeomorfismos. Diremos entonces que un sistema mecánico es simétrico ante un grupo de transformaciones de su variedad de configuración si su lagrangiana es invariante ante el grupo. Naturalmente, una vez que tenemos un sistema de coordenadas generalizadas, podemos decir que el sistema mecánico es simétrico ante un grupo de transformaciones de su variedad de configuración, si la expresión de la lagrangiana en el sistema de coordenadas y la que se obtiene al realizar cualquiera de las transformaciones de coordenadas correspondientes, es la misma, o sea: si $(q^1, \cdots, q^n)$ son unas coordenadas y $(Q^1_\alpha, \cdots, Q^n_\alpha)$ son las coordenadas transformadas, la expresión de la lagrangiana en las primeras, $L(q^1, \cdots q^n, \dot{q}^1, \cdots, \dot{q}^n)$ y en las segundas

$$L\left(q^1(Q_\alpha), \cdots, q^n(Q_\alpha), \dot{q}^1(Q_\alpha, \dot{Q}_\alpha), \cdots, q^n(Q_\alpha, \dot{Q}_\alpha)\right) \tag{3.252}$$

es la misma.

Cuando un sistema mecánico que no depende explícitamente del tiempo presenta simetría ante un grupo de transformaciones de un parámetro de su variedad de configuración, o sea, es invariante ante todas las transformaciones de ese grupo, ocurre que, por ese hecho, tiene una integral de movimiento, veamos esto con detenimiento.

Consideremos unas coordenadas generalizadas para nuestro sistema $(q^1, \cdots, q^n)$ y llamemos α al parámetro del grupo de transformaciones de la variedad de configuración en si misma. Podemos estudiar las transformaciones de la variedad estudiando las transformaciones de coordenadas correspondientes. Llamemos $(Q^1_\alpha, \cdots, Q^n_\alpha)$ a las coordenadas transformadas, de tal suerte que

habrá una serie de relaciones de la forma

$$\begin{aligned} Q^i_\alpha &= Q^i_\alpha(q^1,\cdots,q^n), \\ \dot{Q}^i_\alpha &= \sum_j \frac{\partial Q^i_\alpha}{\partial q^j}\dot{q}^j = \dot{Q}^i_\alpha(q^1,\cdots,q^n,\dot{q}^1,\cdots,\dot{q}^n), \end{aligned} \tag{3.253}$$

o en su expresión inversa

$$q^i = q^i(Q^1_\alpha,\cdots,Q^n_\alpha), \tag{3.254}$$

$$\dot{q}^i = \dot{q}^i(Q^1_\alpha,\cdots,Q^n_\alpha,\dot{Q}^1_\alpha,\cdots,\dot{Q}^n_\alpha). \tag{3.255}$$

La lagrangiana del sistema quedará expresada en términos de las coordenadas transformadas y será una función del parámetro α, podremos calcular su derivada total con respecto a ese parámetro, obtendremos

$$\begin{aligned} \frac{dL}{d\alpha} &= \sum_i \left[\frac{\partial L}{\partial Q^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha} + \frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{\partial \dot{Q}^i_\alpha}{\partial \alpha}\right] \\ &= \sum_i \left[\frac{\partial L}{\partial Q^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha} + \frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{d}{dt}\left(\frac{\partial Q^i_\alpha}{\partial \alpha}\right)\right] \\ &= \sum_i \left[\frac{\partial L}{\partial Q^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha} + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\right)\frac{\partial Q^i_\alpha}{\partial \alpha}\right] \\ &= \sum_i \left[\frac{\partial Q^i_\alpha}{\partial \alpha}\left(\frac{\partial L}{\partial Q^i_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\right)\right) + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha}\right)\right], \end{aligned} \tag{3.256}$$

valuando esta expresión en $\alpha = 0$ y recordando que la transformación para ese valor de α es la transformación idéntica, esto es $Q^i_{\alpha=0} = q^i$, obtenemos

$$\left.\frac{dL}{d\alpha}\right|_{\alpha=0} = \sum_i \left[\left.\frac{\partial Q^i_\alpha}{\partial \alpha}\right|_{\alpha=0}\left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right) + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha}\right)_{\alpha=0}\right]. \tag{3.257}$$

Ahora sobre las trayectorias del sistema se deben cumplir las ecuaciones de Lagrange por lo que los primeros términos del lado derecho de (3.257) son cero y obtenemos

$$\left.\frac{dL}{d\alpha}\right|_{\alpha=0} = \frac{d}{dt}\sum_i \left(\frac{\partial L}{\partial \dot{Q}^i_\alpha}\frac{\partial Q^i_\alpha}{\partial \alpha}\right)_{\alpha=0}; \tag{3.258}$$

si el sistema mecánico es simétrico ante el grupo de transformaciones, la derivada de su lagrangiana con respecto a α debe ser cero y (3.258) implica

la existencia de una integral de movimiento dada por la expresión

$$I = \sum_i \left(\frac{\partial L}{\partial \dot{Q}^i_\alpha} \frac{\partial Q^i_\alpha}{\partial \alpha} \right)_{\alpha=0} = \sum_i \frac{\partial L}{\partial \dot{q}^i} \frac{\partial Q^i_\alpha}{\partial \alpha} \bigg|_{\alpha=0}. \tag{3.259}$$

Este resultado se conoce como el teorema de Noether y nos dice que si un sistema mecánico es simétrico ante un grupo continuo de un parámetro de transformaciones de su variedad de configuración, entonces existe una integral de movimiento dada por la expresión (3.259). Naturalmente si hay una dependencia explicita en el tiempo se obtendrá una constante de movimiento, mas no una integral de movimiento.

La relación entre integrales de movimiento y coordenadas ignorables puede verse como un caso particular de este teorema, veamos. Al existir una variable ignorable, digamos q^i, existe una simetría ante las translaciones a lo largo de esa coordenada. Estas translaciones dejan a todas las coordenadas sin cambios excepto a la iésima que quedará como

$$Q^i_\alpha = q^i + \alpha, \tag{3.260}$$

la velocidad generalizada $\dot{q}^i$ quedará invariante y, como lo único que cambia es la iésima coordenada y ésta no aparece en la lagrangiana, el sistema mecánico es simétrico ante estas transformaciones y habrá una integral de movimiento. Al aplicar (3.259) a este caso obtenemos el resultado esperado

$$I = \frac{\partial L}{\partial \dot{q}^i}. \tag{3.261}$$

Esto es, el impulso generalizado correspondiente a una coordenada será constante si el sistema es constante ante traslaciones a lo largo de esa coordenada.

Veamos ahora el caso del campo central en dos dimensiones, pero formulado en coordenadas cartesianas x y y. La lagrangiana del sistema es

$$L = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) - V(x^2 + y^2). \tag{3.262}$$

Las rotaciones por un ángulo α se expresan como una rotación de las coordenadas, o sea las transformaciones de coordenadas

$$X_\alpha = x \cos\alpha - y \sin\alpha, \tag{3.263}$$

$$Y_\alpha = x \sin\alpha + y \cos\alpha, \tag{3.264}$$

$$\dot{X}_\alpha = \dot{x} \cos\alpha - \dot{y} \sin\alpha, \tag{3.265}$$

$$\dot{Y}_\alpha = \dot{x}\sin\alpha + \dot{y}\cos\alpha. \tag{3.266}$$

El aplicar estas transformaciones en (3.262) nos hace ver la invariancia de la lagrangiana, ya que nos dará

$$L = \frac{1}{2}m(\dot{X}_\alpha^2 + \dot{Y}_\alpha^2) - V(Y_\alpha^2 + Y_\alpha^2), \tag{3.267}$$

por tanto habrá una integral de movimiento asociada a esta simetría del sistema. La fórmula (3.259) nos da para esta integral de movimiento la expresión

$$I = m(\dot{y}x - \dot{x}y), \tag{3.268}$$

que identificamos como el impulso angular. En coordenadas polares las rotaciones se ven como translaciones a lo largo de la coordenada θ que es ignorable y obtenemos el mismo resultado, como era de esperarse.

En caso de que la lagrangiana dependiera explícitamente del tiempo o para incluir transformaciones que afectaran al tiempo, podemos extender lo hecho por medio de la siguiente construcción.

En primer lugar extendemos la variedad de configuración M en una dimensión más construyendo la variedad $M_{\text{ext}} = M \times R$ e introducimos a t como la coordenada generalizada correspondiente a esa dimensión adicional. Pasemos ahora a construir un sistema mecánico en esta variedad de configuración de $n+1$ dimensiones. Comencemos, para no generar confusiones, por denotar al parámetro tiempo de este nuevo sistema por la letra τ y a las derivadas de las coordenadas generalizadas con respecto a τ por una barra esto es

$$\bar{q}^i = \frac{dq^i}{d\tau} \quad y \quad \bar{t} = \frac{dt}{d\tau}. \tag{3.269}$$

Para obtener la lagrangiana del nuevo sistema, extenderemos la lagrangiana del sistema original a M_{ext} por medio de

$$L_{\text{ext}}(q, t, \bar{q}, \bar{t}) = L(q, \frac{\bar{q}}{\bar{t}}, t)\bar{t}. \tag{3.270}$$

Esta L_{ext} no depende explicitamente de τ y dará los mismos resultados que L cuando la coordenada t se reduzca a τ, ya que en ese caso $\bar{t} = 1$, $\bar{q}^i = \dot{q}^i$ y $L_{\text{ext}} = L(q, \dot{q}, t)$.

Ahora podemos contemplar un grupo de un parámetro de transformaciones de M_{ext} en si misma, o sea que, naturalmente puede incluir transformaciones del tiempo

$$Q_\alpha^i = Q_\alpha^i(q, t), \tag{3.271}$$

$$T_\alpha = T_\alpha(q, t). \tag{3.272}$$

El teorema de Noether se aplicará exactamente igual, esto es, si $L_{\rm ext}$ es invariante ante el grupo de transformaciones, entonces existirá una cantidad $I_{\rm ext}$, independiente de τ, tal que

$$\frac{dI_{\rm ext}}{d\tau} = 0, \tag{3.273}$$

con

$$I_{\rm ext} = \sum_i \frac{\partial L_{\rm ext}}{\partial \bar{q}^i} \frac{\partial Q^i_\alpha}{\partial \alpha}\bigg|_{\alpha=0} + \frac{\partial L_{\rm ext}}{\partial \bar{t}} \frac{\partial T_\alpha}{\partial \alpha}\bigg|_{\alpha=0}, \tag{3.274}$$

que es la expresión (3.259) para este caso. Vemos que si $\bar{t} = 1$ entonces $I_{\rm ext}(q,t,\bar{q},\bar{t}) = I(q,\dot{q},t)$ y esta I será tal que[14]

$$\frac{dI}{dt} = 0. \tag{3.275}$$

En particular si L no depende explícitamente del tiempo, t será una coordenada ignorable de $L_{\rm ext}$ y el sistema será invariante ante el grupo de translaciones en el tiempo

$$Q^i_\alpha = q^i, \tag{3.276}$$
$$T_\alpha = t + \alpha, \tag{3.277}$$

existirá, por tanto, una cantidad constante $I_{\rm ext}$. En este caso, por ser L independiente del tiempo, al hacer $\tau = t$ está cantidad se reducirá a una integral de movimiento del sistema original. Al hacer las substituciones correspondientes en (3.274) obtenemos

$$I = -\sum_i \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i + L(q,\dot{q}), \tag{3.278}$$

que es la negativa de la expresión (3.125), que ya habíamos identificado con una integral de movimiento cuando la lagrangiana no depende explícitamente del tiempo. Aquí vemos el origen de esta integral de movimiento, la invariancia del sistema mecánico ante el grupo de translaciones en el tiempo. Recordamos también que en los caso en que la energía cinética es una forma cuadrática en las velocidades generalizadas, esta integral de movimiento es la energía total, por lo que podemos decir que la conservación de la energía en estos sistemas mecánicos, tiene origen en su invariancia ante translaciones en el tiempo. Además podemos observar de (3.274) que la integral de

[14] En general esta I no será una integral de movimiento del sistema original debido a su dependencia explícita del tiempo, será simplemente una constante de movimiento.

movimiento o, en su caso la energía total, es la derivada parcial de L_{ext} con respecto a $\bar{t}$, o sea, podemos decir que la energía total es el impulso generalizado correspondiente a la coordenada generalizada tiempo y que cuando el tiempo es ignorable ese impulso generalizado es una integral de movimiento; esto es similar a lo que sucede con cualquier otra coordenada generalizada y su correspondiente impulso generalizado. Esta relación entre el tiempo y la energía total se extiende a otras teorías físicas, como la mecánica cuántica, y es la razón para que el principio de incertidumbre se aplique cuando se miden simultáneamente el tiempo y la energía de un sistema cuántico.

El teorema de Noether tiene gran importancia para el pensamiento físico, pues, al relacionar las simetrías con las integrales de movimiento, hace ver claramente la relación entre la situación geométrica y dinámica de un sistema y sus propiedades físicas.

En el contexto de la mecánica clásicas, sin embargo, no resulta de mucha utilidad práctica debido a que en general es tan difícil encontrar las simetrías que un sistema pudiera tener, como encontrar las integrales de movimiento correspondientes por otros procedimientos. Esto no es así cuando abordamos otros campos de la física a los que el teorema de Noether se extiende de manera natural, en especial en el caso de la mecánica cuántica y la teoría de los campos; en esos casos estamos interesados muchas veces en construir teorías que expliquen alguna propiedad física que conocemos a partir de experimentos y que sabemos se corresponde a una integral de movimiento en el caso clásico, y por tanto con alguna simetría; esto nos permite restringir la búsqueda de lagrangianas, o sus equivalentes en el contexto en que estemos, a aquellas que presenten dicha simetría y por esta vía simplificar enormemente la tarea de encontrar teorías satisfactorias.

3.10. Variedades diferenciales.

A lo largo de este capítulo hemos mencionado en diversas ocasiones que los espacios en que representamos los sistemas físicos no son R^n, es más no son siquiera topológicamente equivalentes a R^n. Esto introduce una serie de dificultades, tanto de carácter práctico, como conceptuales. Hemos dicho que estas dificultades se resuelven en el contexto del cálculo sobre variedades diferenciales, pero nos hemos limitado a señalar que podemos trabajar como si estuviéramos en R^n, siempre y cuando consideremos que nuestros resultados son únicamente locales. Con el propósito de dejar esto claro, a continuación se dan los elementos más importantes de la teoría de las variedades diferenciales y del cálculo sobre ellas.

Comencemos por establecer las dificultades que aparecen al pretender extender el cálculo diferencial, digamos, a la superficie de una esfera.

El cálculo tiene por objetivo el estudio de los mapeos entre espacios, así el llamado cálculo de varias funciones de varias variables es el estudio de los mapeos de R^n en R^m. Un primer punto a señalar es que los dos espacios deben ser espacios topológicos para poder introducir el concepto de continuidad de los mapeos. En el cálculo común para los espacios R^n se toma la topología inducida por la métrica euclidiana. Para el objetivo de esta sección supondremos que siempre partimos de espacios topológicos, esto es, está bien definido lo que es un subconjunto abierto de los puntos que componen el espacio y estos subconjuntos forman una topología[15].

La idea básica del cálculo diferencial consiste en estudiar los mapeos entre espacios aproximándolos, en cada punto, por un mapeo lineal, naturalmente se utiliza, en cada punto, aquel mapeo lineal que mejor aproxima al mapeo en ese punto, ese mapeo lineal es la diferencial del mapeo en dicho punto. Veamos como ejemplo el caso más simple un mapeo o función de R en R, $y = f(x)$. Estamos acostumbrados a ver la derivada de f valuada en un punto como un número, sin embargo ese número representa una función lineal de R en R, o sea es la matriz de uno por uno que, en las coordenadas utilizadas, expresa a la función lineal. En forma geométrica representamos a la función por su gráfica en coordenadas (x, y) y a su derivada por la tangente a esa gráfica en el punto de que se trate; ahora este espacio (x, y) sólo tiene sentido para trazar gráficas de mapeos de R en R, por tanto la linea tangente se debe interpretar como la gráfica de un mapeo que, por ser una recta, es la gráfica de un mapeo lineal y que, por ser tangente a la gráfica de la función en el punto de interés, es la mejor aproximación lineal a $f(x)$ en ese punto.

Ahora para poder hablar de mapeos lineales, los espacios entre los que se establecen dichos mapeos deben ser espacios vectoriales, de lo contrario no tiene sentido hablar de linealidad. R^n es un espacio vectorial, por eso podemos construir el cálculo basándonos en él, pero los espacios que se generan en la mecánica, como la esfera que es, por ejemplo, el espacio de configuración del péndulo esférico, si bien son espacios topológicos, en general no pueden ser espacios vectoriales, para ver esto basta recordar que todo espacio vec-

[15] Dado un conjunto no vacío una topología es una colección de subconjuntos de él tales que: el vacío forme parte de la colección, el conjunto mismo forme parte de la colección, la unión de cualquier número de subconjuntos de la colección pertenezca a la colección y la intersección de un número finito de subconjuntos de la colección pertenezca a la colección. A un conjunto al que se le ha asociado una topología se le llama espacio topológico y a los subconjuntos de la colección que conforman su topología se les llama subconjuntos abiertos o simplemente abiertos.

torial es homeomorfo a R^n para alguna n y, en general, los espacios de la mecánica, como la esfera, no lo son. Así la principal dificultad que surge al tratar con mapeos entre espacios, que no son espacios vectoriales, consiste en que no podemos definir la diferenciabilidad, ya que, para hacerlo requerimos del concepto de linealidad.[16]

Tomemos dos espacios topológicos T_1 y T_2 y un mapeo entre ellos $f : T_1 \longrightarrow T_2$. Supongamos que T_1 es homeomorfo a R^n, esto es existe un homeomorfismo h_1 entre ellos, y que T_2 es homeomorfo a R^m, denotemos este último homeomorfismo por h_2. Entonces el siguiente diagrama

$$\begin{array}{ccc} & f & \\ T_1 & \longrightarrow & T_2 \\ h_1 \updownarrow & & \updownarrow h_2 \\ \\ R^n & \longrightarrow & R^m \qquad (3.279) \\ & h_2 \circ f \circ h_1^{-1} & \qquad (3.280) \end{array}$$

nos hace ver que podemos inducir un mapeo de R^n en R^m usando los homeomorfismos y el mapeo f, como ahora tenemos una mapeo que refleja al mapeo f pero que se establece entre espacios vectoriales, podemos hacer la definición que requerimos y después, usando nuevamente los homeomorfismos, trasladar los resultados al mapeo original. De hecho esto es lo que hacemos normalmente al estudiar el cálculo diferencial en coordenadas curvilíneas, ya que, en vez de pensar en un espacio R^n euclidiano en el que hemos introducido un sistema de coordenadas curvilíneas, podemos pensar en que el espacio está curvado y que la transformación de coordenadas es un homeomorfismo entre ese espacio curvado y R^n con su estructura euclidiana, con lo cual llegamos al mismo diagrama, procedimiento y resultado.

Este procedimiento no es válido cuando los espacios topológicos no son homeomorfos a R^n, como ocurre con la esfera que hemos tomamos como ejemplo ilustrativo. Sin embargo puede ocurrir que algunos de los subconjuntos abiertos de nuestros espacios topológicos si sean homeomorfos a subconjuntos abiertos de R^n, por ejemplo cualquier pedazo de la esfera lo es, de hecho toda la esfera menos un punto es un subconjunto abierto de ella homeomorfo a R^2. Puede ocurrir además que tengamos varios subconjuntos abiertos homeomorfos a R^n y que entre todos cubran al espacio topológico

[16] Dos espacios topológicos son homeomorfos si existe entre ellos un mapeo biunívoco y continuo en ambas direcciones, a un mapeo así se le conoce como homeomorfismo, esta noción es la misma que la de equivalencia topológica.

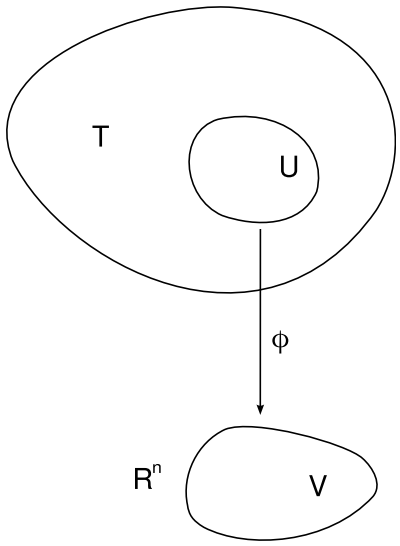


Figura 3.14: Una carta es un homeomorfismo ϕ entre un abierto U del espacio topológico T y un abierto V de R^n.

completo, esto es que su unión sea todo el espacio. En ese caso podemos establecer el procedimiento anterior para cada abierto y si logramos que en los traslapes o intersecciones entre abiertos el procedimiento usando uno de ellos y su correspondiente homeomorfismo y el procedimiento usando el otro abierto y su homeomorfismo sean congruentes, entonces podremos extender el cálculo a esos espacios. Esto es, podremos calcular por pedazos, sabiendo que en los traslapes los cálculos resultan consistentes y que entre todos los pedazos se cubre la totalidad del espacio. Esta idea es la base de la definición de variedad diferencial y de la construcción del cálculo diferencial para mapeos entre esas variedades. Pasemos pues a hacer las definiciones y deducciones correspondientes. Acompañaremos esto de una serie de figuras que, utilizando la esfera como ejemplo, ilustran el proceso.

Tomemos un espacio topológico T, una carta consiste de un subconjunto abierto de T, $U \subset T$ y un homeomorfismo ϕ entre U y un subconjunto abierto de R^n, $V \subset R^n$. Nótese que pueden existir espacios topológicos en los que no se puede definir ninguna carta, esto ocurre cuando ninguno de sus abiertos es homeomorfo a R^n, esos están excluidos de esta construcción.

Ahora consideremos dos cartas (U_1, ϕ_1, V_1) y (U_2, ϕ_2, V_2) para las que los abiertos U_1 y U_2 se intersectan. Para los puntos comunes de U_1 y U_2 tendremos dos homeomorfismos distintos con dos subconjuntos distintos de

R^n.[17] La intersección se proyectará en dos 'pedazos' de R^n, uno en cada uno de los dos subconjuntos de R^n y podremos establecer un mapeo entre estos dos 'pedazos' por medio de siguiente procedimiento: tomemos un punto de V_1 correspondiente a un punto de la intersección $U_1 \cap U_2$, usando el primer homeomorfismo ϕ_1^{-1} lo mandamos a su imagen en la intersección, como está en la intersección estará también en U_2 y se puede mapear a V_2 por medio de ϕ_2, esto es construimos el mapeo $\phi_2 \circ \phi_1^{-1}$ de la imagen de la intersección $U_1 \cap U_2$ en V_1 generada por ϕ_1 a la imagen de la misma intersección en V_2 generada por ϕ_2. Como estas imágenes son subconjuntos abiertos de R^n podemos utilizar todos los conceptos del cálculo para los mapeos entre ellas, diremos, entonces, que las dos cartas son compatibles si el mapeo de R^n en R^n así definido es un difeomorfismo, esto es si es biunívoco y es diferenciable en ambos sentidos.[18] Diremos, por extensión, que dos cartas son compatibles si los abiertos en que se sustentan son ajenos.

Un atlas es una colección de cartas compatibles entre si que cubren todo el espacio topológico, esto es, tal que la unión de todos los abierto de la colección de cartas sea todo el espacio.

Finalmente definimos una variedad diferencial como un espacio topológico dotado de un atlas, se dice que el atlas define una estructura diferencial para la variedad.[19] Diremos que la variedad diferencial tiene la dimensión que tenga el espacio cartesiano en el que se establecen las cartas.

Hablando coloquialmente una variedad diferencial es un espacio topológico M hecho de parches, en el que cada parche es esencialmente un abierto de R^n. Estos parches son abiertos de M, preservan su estructura topológica y se pegan consistentemente en el sentido diferencial.

Como ejemplo tomemos la esfera de dos dimensiones. La esfera no es homeomorfa a R^2, pero el abierto que resulta de quitarle un punto si lo es, por tanto podemos escoger un punto, quitarlo y tomar el abierto que

[17] Se puede demostrar fácilmente que si un abierto de un espacio topológico es homeomorfo a un abierto de R^n, entonces cualquier otro abierto de ese espacio que sea homeomorfo a un abierto de un espacio cartesiano lo será a R^n, esto es, la dimensión es la misma, de lo contrario se podría establecer un homeomorfismo entre dos espacios cartesianos de diferente dimensión.

[18] Se puede pedir que el mapeo y su inversa sean de clase C^k, en cuyo caso se dice que las dos cartas son compatibles de clase C^k. En este texto siempre suponemos que los mapeos con que trabajamos son suficientemente derivables para nuestros propósitos. En la mecánica clásica suele ser suficiente con que sean de clase C^2, pero para no distraer la atención con estos detalles hemos pedido que sean diferenciables, el lector interesado podrá hacer los ajustes necesarios.

[19] En rigor debemos ampliar el atlas a un atlas maximal, esto es agregar al atlas todas las posibles cartas compatibles con él

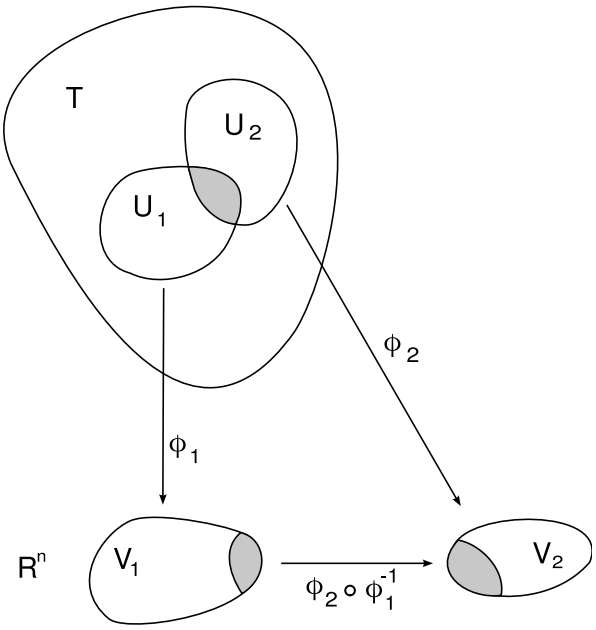


Figura 3.15: Compatibilidad de dos cartas.

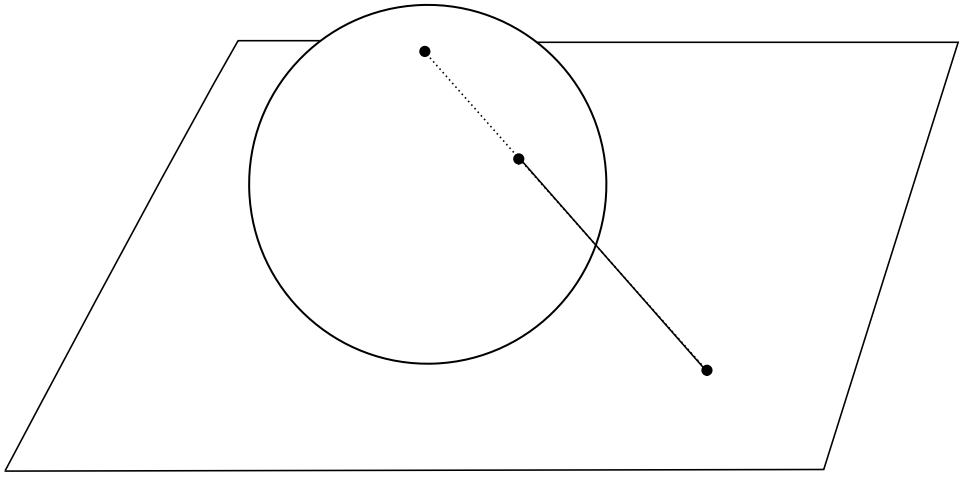


Figura 3.16: Proyección estereográfica de la esfera en un plano. Esta proyección forma una carta para el abierto consistente en la esfera con la exclusión del polo superior.

resulta para construir una carta. Para hacer esto tenemos que especificar un homeomorfismo entre ese abierto y R^2. Tomemos como punto a excluir, por ejemplo, el "polo superior", como se muestra en la figura (3.16), ahora tomemos un plano, esto es R^2, coloquémoslo tangente a la esfera en el "polo inferior" y generemos la proyección de los puntos de la esfera, excluido el "polo superior", al plano por medio de la recta que parte del "polo superior" y pasa por el punto en cuestión. Evidentemente se trata de una relación biunívoca y continua en ambas direcciones, esto es un homeomorfismo, por tanto ya tenemos una carta, en cierto sentido es una buena carta ya que abarca gran parte de la esfera (toda menos un punto). Podemos ahora construir otra carta por un proceso similar, pero quitando ahora el "polo inferior", colocando al plano tangente a la esfera en el "polo superior" y haciendo la proyección a partir del "polo inferior". Estas dos cartas cubren toda la esfera y es fácil hacer ver que el mapeo que se establece de R^2 en R^2 en donde hay traslape, o sea en toda la esfera menos los dos "polos", es un difeomorfismo, para hacer esto podemos poner coordenadas cartesianas en el plano de "abajo" y en el de "arriba" y expresar en esas coordenadas el mapeo compuesto entre las dos proyecciones.

El resultado es

$$\begin{aligned} x_2 &= 4r^2 \frac{x_1}{x_1^2 + y_1^2}, \\ y_2 &= 4r^2 \frac{y_1}{x_1^2 + y_1^2}, \end{aligned} \tag{3.281}$$

donde r es el radio de la esfera. Vemos que el mapeo es biunívoco y diferenciable excepto en $(x_1, y_1) = (0, 0)$ y en $(x_2, y_2) = (0, 0)$, pero esos puntos no están en la intersección de las dos cartas.

Claro que podemos construir muchas otras cartas e incluso muchos otros atlas, pero, en el caso de la esfera de dos dimensiones, encontraríamos que todas las cartas resultarían compatibles con las dos descritas aquí. Esto implica que solo existe una estructura diferencial para la esfera de dos dimensiones, o sea con la topología de una esfera de dos dimensiones solo hay una variedad diferencial. Milnor demostró que, en el caso de la esfera de siete dimensiones, se pueden construir atlas distintos, esto es, atlas tales que las cartas de uno son incompatibles con las del otro; así para siete dimensiones hay varias variedades diferenciales distintas con la topología de la esfera; nótese que un mapeo que involucre a la esfera de siete dimensiones puede ser diferenciable en un atlas y no serlo en otro, esto es tener ciertas propiedades en uno y otras en otro, o sea en realidad se trata de variedades diferenciales distintas, aunque topológicamente sean el mismo espacio. Esto ocurre en general, esto

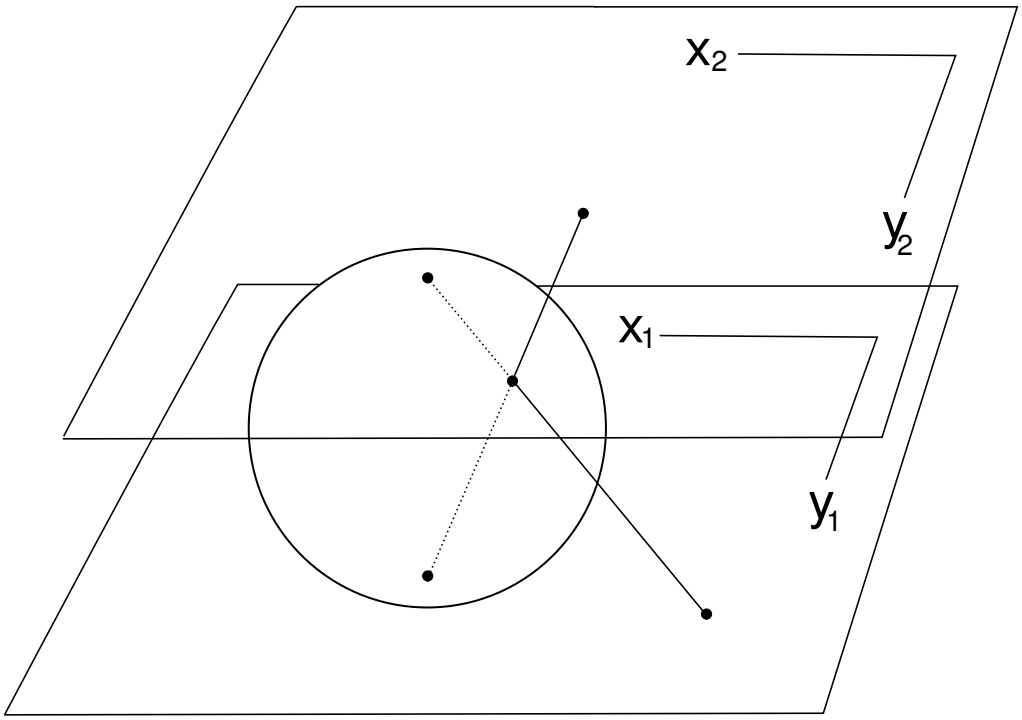


Figura 3.17: Un atlas para la esfera.

es, sustentados en un espacio topológico específico, pueden construirse varias estructuras diferenciales distintas y, por tanto, varias variedades diferenciales distintas.

Para tratar problemas concretos se requiere introducir coordenadas, en el caso de las variedades diferenciales se hace introduciendo coordenadas locales, veamos esto. Cuando estamos en una variedad diferencial tomamos una carta, recordemos que una carta es un abierto y un homeomorfismo de ese abierto a R^n, en R^n ponemos un sistema de coordenadas, el que nos sea de mayor utilidad; entonces usando el homeomorfismo le hacemos corresponder a puntos de la variedad puntos de R^n, los cuales quedan descritos por las coordenadas de R^n que les correspondan, podemos así decir que hemos puesto coordenadas a los puntos de la variedad. Claro está que, como la carta no cubre toda la variedad, solo tendrán coordenadas los puntos del abierto correspondiente, por eso llamamos a estas coordenadas, coordenadas locales. Cuando estamos interesados en puntos de la variedad que no han sido cubiertos por esas coordenadas locales, debemos buscar otra carta del atlas que si los contenga y construir otro sistema de coordenadas locales apropiado para esos puntos y así sucesivamente. La compatibilidad entre las cartas nos garantiza que al hacer los pasos de un sistema local de coordenadas a otro, se conservan las propiedades de diferenciabilidad. En muchas ocasiones, como en la esfera de dos dimensiones que hemos tomado como ejemplo, se pueden tomar sistemas locales de coordenadas que cubran una gran parte de la variedad, en nuestro ejemplo toda la variedad excepto un punto, pero si una variedad no es homeomorfa a R^n se requerirá de varias cartas para cubrirla, pues no existe ningún sistema local de coordenadas que la cubra totalmente.

Las coordenadas generalizadas de la mecánica, no son otra cosa que sistemas locales de coordenadas para la variedad de configuración. Ocurre que en muchas ocasiones podemos cubrir gran parte de esa variedad con solo un sistema local de coordenadas, pero, como hemos dicho, siempre hay que tener en cuenta que algo estamos dejando fuera y que pueden presentarse casos en que se dejan fuera cuestiones importantes. Adicionalmente hay que señalar que siempre que consideramos un solo sistema local de coordenadas, por mucho que cubra, además de dejar fuera algunos puntos de la variedad, dejamos fuera la estructura topológica de esta, cuestión que puede conducirnos a errores importantes.

Veamos ahora como podemos extender algunas cuestiones básicas del cálculo diferencial en R^n a las variedades diferenciales.

Supongamos que tenemos dos variedades diferenciales N y M y un mapeo

entre ellas

$$F : N \longrightarrow M. \tag{3.282}$$

Consideremos un punto arbitrario $x \in N$ y una carta de N que lo cubra (U_1^N, ϕ_1^N). Tomemos ahora una carta de M, (U_1^M, ϕ_1^M), que cubra a $F(x)$, la imagen de x en M. Por medio del diagrama

$$\begin{array}{ccc} & F & \\ U_1^N \subset N & \longrightarrow & U_1^M \subset M \\ \phi_1^N \updownarrow & & \updownarrow \phi_1^M \\ V_1^N \subset R^n & \longrightarrow & V_1^M \subset R^m \\ & f & \end{array} \tag{3.283}$$

generamos un mapeo f de R^n en R^m, dado por $f = \phi_1^M \circ F \circ \left(\phi_1^N\right)^{-1}$. Diremos que el mapeo F es diferenciable en el punto x si el mapeo f de R^n en R^m es diferenciable en el punto $\phi_1^N(x)$, esto último corresponde al cálculo común donde tiene pleno sentido el concepto de diferenciabilidad. Diremos que el mapeo es diferenciable si es diferenciable en todos los puntos de la variedad.

Ahora bien para que esta definición de diferenciabilidad tenga sentido hace falta que no dependa de las cartas particulares que hemos escogido, esto es que sea una propiedad del mapeo y no de las cartas. Para hacer ver que esto es así tomamos una segunda carta de N, (U_2^N, ϕ_2^N) que contenga a x y otra carta de M, (U_2^M, ϕ_2^M) que contenga a $F(x)$; debemos demostrar que si F es diferenciable en x usando las primeras cartas, también lo será usando las segundas.

En la segunda carta el mapeo generado de R^n en R^m será $f_2 = \phi_2^M \circ F \circ \left(\phi_2^N\right)^{-1}$. Como x está en la intersección de las dos cartas de N y $F(x)$ está en la intersección de las dos cartas de M, por ser N y M variedades diferenciales, se generan dos difeomorfismos, uno de R^n en R^n dado por $\phi_2^N \circ \left(\phi_1^N\right)^{-1}$ o en sentido inverso por $\phi_1^N \circ \left(\phi_2^N\right)^{-1}$ y el otro de R^m en R^m dado por $\phi_2^M \circ \left(\phi_1^M\right)^{-1}$. Ahora tendremos que

$$\phi_2^M \circ f \circ \left(\phi_2^N\right)^{-1} = \left[\phi_2^M \circ \left(\phi_1^M\right)^{-1}\right] \circ \left[\phi_1^M \circ f \circ \left(\phi_1^N\right)^{-1}\right] \circ \left[\phi_1^N \circ \left(\phi_2^N\right)^{-1}\right], \tag{3.284}$$

o sea, el mapeo generado usando las segundas cartas es la composición del generado por las primeras con los difeomorfismos que provienen de la definición de variedad diferencial y por tanto si uno es diferenciable también lo será el otro. Concluimos que el concepto de diferenciabilidad de un mapeo entre variedades diferenciales está bien definido, pues independiente de

las cartas que se usen. Adicionalmente podemos ver que cuando las variedades diferenciales son espacios cartesianos, el concepto de diferenciabilidad se reduce a ese concepto en el marco del cálculo diferencial común.

En general cuando trabajamos con variedades diferenciales y generamos definiciones, cálculos o teoremas y sus demostraciones, debemos seguir procedimientos semejantes al que acabamos de recorrer para demostrar que lo que hacemos es independiente de las cartas o coordenadas locales que utilicemos; de lo contrario no estaremos obteniendo propiedades de las variedades diferenciales y sus mapeos, sino propiedades de las cartas o coordenadas locales que hayamos escogido y lo que queremos estudiar son las variedades y sus mapeos y no la arbitrariedad que introducimos al escoger una carta en particular.

Hemos establecido el concepto de diferenciabilidad de los mapeos entre variedades diferenciales, pensaríamos en pasar a definir el concepto de diferencial para esos mapeos. Para hacerlo podríamos tomar el mapeo entre R^n y R^m inducido por una carta y definir la diferencial del mapeo entre las variedades como la diferencial de ese mapeo inducido, sin embargo cometeríamos por este camino un grave error: contrariamente a lo que ocurre con la diferenciabilidad, la diferencial dependería de las cartas que hayamos escogido, lo que arruina totalmente nuestro intento. Hay, desde luego, una forma correcta de definir la diferencial, pero para poder hacerlo debemos pasar antes a estudiar algunas otras cuestiones, las que, adicionalmente, son de importancia para la mecánica clásica y mucha otras disciplinas.

En un sistema mecánico las trayectorias, o sea las curvas que traza el punto representativo del sistema en la variedad de configuración como resultado del movimiento, son de la mayor importancia, en este sentido el estudio de las curvas sobre variedades diferenciales adquiere relevancia.

Una curva sobre una variedad es un mapeo, que supondremos diferenciable, de R en la variedad. En general usaremos la letra t para denotar la coordenada en R, o, dicho de otra forma, para parametrizar la curva. Ahora bien, en la mecánica nos interesa tanto la curva o trayectoria como la velocidad con que ésta se recorre. La velocidad debe ser un vector, pero la curva está sobre una variedad diferencial que no es un espacio vectorial, por lo que no puede ser un vector de la variedad, sin embargo no requerimos más que R, la variedad y el mapeo para que la curva esté bien definida. Tenemos pues un problema: a que espacio vectorial pertenece el vector velocidad.

Cuando contemplamos el ejemplo del péndulo esférico no podemos olvidar que la esfera está inmersa en el espacio físico de tres dimensiones que si es lineal, por tanto podemos pensar que la velocidad es un vector en el espacio físico; pero si centramos nuestra atención en el caso del péndulo doble las

cosas cambian, la variedad de configuración es un toro de dos dimensiones cuyos puntos representan la posición de las dos partículas que componen el péndulo doble, este toro no está inmerso en el espacio físico, es más no está inmerso en ningún espacio, por lo que no queda claro a que espacio vectorial pertenece entonces la 'velocidad'.

Naturalmente toda variedad diferencial puede sumergirse en un espacio lineal[20], pero no tiene sentido el construir un espacio artificial únicamente para poder ver la variedad de configuración inmersa en él, cuando es mucho mejor concebirla por si misma. Es más, hay un asunto profundamente físico que nos debe hacer ver que la velocidad no es un vector del espacio físico, ni siquiera en los casos en que la trayectoria está en el espacio físico, como sería el caso en un tiro parabólico; simplemente en el espacio físico las dimensiones son de distancia y el vector velocidad tiene dimensiones de velocidad, por tanto no puede ser un vector del espacio físico.

Para resolver esta situación debemos desarrollar algunos conceptos en torno de las variedades diferenciales.

Consideremos una variedad diferencial N de dimensión n y un punto x de ella. Por ese punto pasa una inmensidad de curvas, todos los mapeos de R en la variedad que para algún valor de t den x. Todas estas curvas pueden ser agrupadas en clases de la siguiente manera. Diremos que dos de esas curvas son tangentes en el punto x, si al tomar una carta de la variedad que cubra a x y proyectar, por medio del homeomorfismo de la carta, las curvas resultantes en R^n son tangentes en el punto $\phi(x)$. Debemos, claro está, demostrar que esta definición es independiente de la carta que tomemos, esto es que si dos curvas resultan tangentes en una carta, lo serán en cualquier otra.

Tomemos primero una carta (U_1, ϕ_1) que cubra a x y dos curvas c_1 y c_2 tales que $c_1(t_1) = c_2(t_2) = x$ y que en esa carta sean tangentes, esto es, las dos curvas proyectadas a R^n pasan por el punto $\phi_1(x)$ y

$$\left.\frac{d}{dt}\left(\phi_1 \circ c_1\right)\right|_{t_1} = \left.\frac{d}{dt}\left(\phi_1 \circ c_2\right)\right|_{t_2}. \tag{3.285}$$

Tomemos ahora otra carta (U_2, ϕ_2), se establece entonces un difeomorfismo de R^n, $\phi_1 \circ \phi_2^{-1}$ o en sentido inverso $\phi_2 \circ \phi_1^{-1}$; en adición a las proyecciones por medio de ϕ_1 tendremos las proyecciones de las curvas a R^n por medio del homeomorfismo ϕ_2. Estas nuevas proyecciones son curvas distintas, pasarán ambas por el punto $\phi_2(x)$ que es distinto a $\phi_1(x)$ y los valores de sus

[20]Este espacio no necesariamente tiene una dimensión más, como ocurre con la esfera, por ejemplo la botella de Klein es una variedad no orientable de dimensión dos, que no se puede sumergir en un espacio de tres dimensiones, requerimos para ella al menos cuatro.

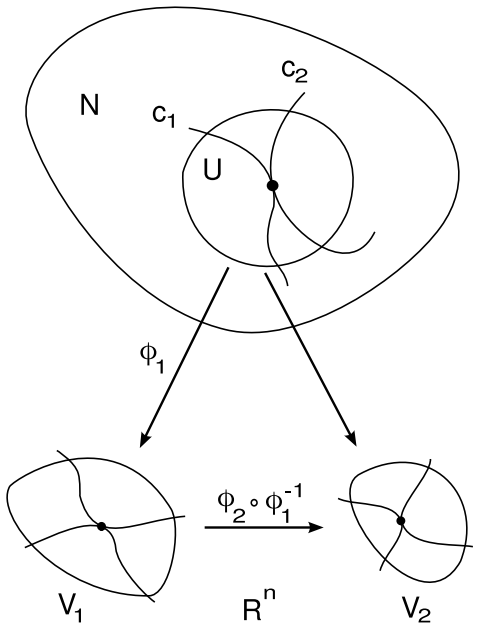


Figura 3.18: Curvas tangentes se mapean en curvas tangentes independientemente de la carta.

derivadas con respecto al tiempo serán distintas a los expresados en (3.285), pero se transformarán en las obtenidas por la primera carta por medio de un difeomorfismo de R^n y por tanto se conservará el que sean o no sean tangentes puesto que los difeomorfismos en R^n preservan las tangencias. Con esto hemos hecho ver que la definición de tangencia entre dos curvas es válida, pues no depende de la carta que utilicemos. Esto se ilustra en la figura 3.18

Es fácil ver que la relación que se establece entre las curvas que pasan por el punto x por medio de la tangencia es una relación de equivalencia, por tanto el conjunto de todas las curvas que pasan por el punto x queda dividido en clases de equivalencia de curvas tangentes. Denotaremos por $[c]$ a la clase de equivalencia de curvas tangentes en x de la que c es una curva representativa, que, desde luego, puede ser cualquiera de ellas.

Este conjunto de clases puede convertirse en un espacio vectorial. Para ello debemos definir la suma de dos de sus elementos y el producto de uno de ellos por un escalar real. Tomemos dos clases $[c_1]$ y $[c_2]$ y una carta, usando las curvas representativas y proyectándolas en R^n por medio del homeomorfismo de la carta, obtendremos dos curvas en R^n. Estas curvas no serán tangentes y se intersectarán en $\phi(x)$, la proyección del punto x en

R^n. En ese punto las curvas tendrán sus vectores tangentes, recordamos que estamos en R^n, ahora los sumamos y obtenemos un nuevo vector en R^n. Construyamos una tercera curva en R^n que pase por $\phi(x)$ y que tenga al vector suma como vector tangente en $\phi(x)$. Utilicemos ϕ^{-1} para "subir" esa tercera curva a la variedad diferencial y llamemos a la curva resultante c_3, está claro que esta curva pasa por el punto x. Definimos ahora la suma de las clase de equivalencia como la clase de equivalencia de c_3 , esto es

$$[c_1] + [c_2] = [c_3]. \tag{3.286}$$

Naturalmente tendríamos que demostrar que esto no depende de la carta en que se hizo la construcción, ni de las curvas representativas que hayamos elegido. No lo haremos aquí, simplemente señalaremos que es posible hacerlo y que se utilizan razonamientos parecidos a los que ya hemos utilizado en asuntos similares. Para definir el producto por un escalar, o sea por un número real, se hace algo similar. Finalmente se puede hacer ver que con estas definiciones este conjunto de clases de equivalencia es un espacio vectorial de la misma dimensión que la variedad. Los detalles de las demostraciones se pueden encontrar en los textos especializados y se basan en trasladar las propiedades de los vectores tangentes en R^n a las clases de equivalencia utilizando curvas representativas y una carta, posteriormente se hace ver que la demostración es independiente de la carta y de las curvas representativas. Al espacio vectorial así formado se le llama el espacio tangente a la variedad en el punto x y se denota por TN_x.

Si pensamos en la variedad como inmersa en un espacio euclidiano, por ejemplo la esfera en R^3, podemos visualizar el espacio tangente en un punto de la variedad, como un espacio R^n de la dimensión de la variedad y tangente a ésta en el punto, en el caso de la esfera tendríamos un plano tangente, ver la figura 3.19.

Esto es más que una imagen pues podemos ver una curva sobre la variedad como una curva en el espacio euclidiano en que está inmersa la variedad, en ese espacio la curva tiene una tangente en el sentido común y será un vector, si hacemos pasar muchas curvas distintas por ese punto, todas ellas contenidas en la variedad, pero tratadas como curvas en el espacio euclidiano, vemos que los vectores tangentes a todas ellas en el punto forman un espacio vectorial de la dimensión de la variedad y tangente a ésta, este espacio es TN_x. De hecho se puede formular toda la teoría de las variedades diferenciales a partir de concebirlas siempre inmersas en un espacio cartesiano, sin embargo es mejor no hacerlo, pues no es necesario ese espacio cartesiano. En cierta forma al definir la variedad de configuración en base a las ecuaciones que definen a las constricciones holonómicas, la entendimos como una

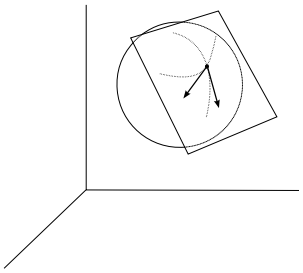


Figura 3.19: El plano tangente en un punto de una esfera y dos de sus vectores vistos desde R^3.

variedad inmersa en el espacio euclidiano que representa al conjunto de las partículas del sistema mecánico, cada una de ellas en el espacio físico de tres dimensiones, el espacio que llamamos espacio de configuración.

Para un sistema mecánico tenemos la variedad de configuración N y curvas sobre ella que son las trayectorias, tenemos además el tiempo, que suponemos es una variable física que transcurre y por tanto la trayectoria se recorre con el transcurrir del tiempo reflejando el movimiento físico. Para poder estudiar esto utilizamos un sistema local de coordenadas, esto es una carta y un sistema de coordenadas en R^n, claro esto, insistimos, es únicamente local, por lo que para estudiar todo el movimiento debemos pasar, cuando esto se requiera, de un sistema local a otro usando la estructura diferencial de la variedad. El sistema local de coordenadas está mediado por el homeomorfismo de la carta, pero es un sistema de coordenadas en R^n, y lo denotamos por $(q^1, \cdots, q^n)$, esto es, son las coordenadas generalizadas y la trayectoria quedará expresada, en esa localidad, como $(q^1(t), \cdots, q^n(t))$[21].

Ahora la 'velocidad' en un cierto punto $q \in N$ de la trayectoria es un vector del espacio tangente a la variedad en ese punto y para poder expresarlo requerimos de una base para el espacio tangente. Podemos, desde luego, tomar cualquier base, pero es sumamente conveniente construir una base que esté relacionada con el sistema de coordenadas locales que estemos utilizando.

Dado un punto $q \in N$ y un sistema local de coordenadas, podemos trazar en R^n las n curvas coordenadas que pasan por el punto $\phi(q)$, las que serán transversales entre si. Ahora podemos, usando el homeomorfismo a la inversa, "subir" esas curvas coordenadas a la variedad, en donde tendremos

[21] Usaremos en lo que sigue la nomenclatura y notación de la mecánica, pero, desde luego, todo esto es totalmente aplicable a las variedades diferenciales en general. En ese caso en vez de hablar de vectores velocidad hablaríamos de vectores tangentes.

n curvas que pasan por el punto q y que serán transversales entre si, por lo que pertenecen a clases de tangencia distintas e independientes; consideradas estas clases como vectores del espacio tangente en q, formarán una base de ese espacio que estará relacionada a las coordenadas locales, denotemos a esta base por $(e_1, \cdots, e_n)$. Para expresar un vector tangente en esta base debemos calcular sus componentes. Una curva representativa de la clase de tangencia en que estemos interesados proyectada en R^n se podrá escribir como $(q^1(t), \cdots, q^n(t))$, por tanto la velocidad o vector tangente en R^n será $(\dot{q}^1(t), \cdots, \dot{q}^n(t))$, pero esto es la representación de la clase de tangencia en R^n y por tanto son las componentes del vector tangente en el sistema local de coordenadas y en la base inducida por esas coordenadas, esto es, lo que en la mecánica llamamos velocidades generalizadas resultan ser las componentes del vector 'velocidad' en la base $(e_1, \cdots, e_n)$ del espacio tangente a la variedad de configuración. Es importante señalar que estos vectores base son distintos cuando se usan distintas coordenadas locales, ya sea por usar una carta distinta, por usar diferentes coordenadas en R^n pero la misma carta o por cambiar ambas.

Nótese que en caso de que la variedad sea R^n mismo, digamos R^3, estos vectores base del espacio tangente son los famosos vectores $\hat{i}$, $\hat{j}$ y $\hat{k}$, si hemos escogido coordenadas cartesianas, $\hat{r}$, $\hat{\theta}$ y $\hat{\varphi}$ si hemos escogido coordenadas esféricas o los que correspondan en cualquier otro sistema de coordenadas; por eso cambian con el punto en que estemos: son vectores base para espacios distintos, un espacio por cada punto, el espacio tangente en ese punto. Naturalmente cuando la variedad es R^n el espacio tangente en un punto se "confunde" con la misma variedad, el hecho de que en coordenadas curvilíneas los vectores de la base cambien con el punto nos recuerda el hecho de que, aunque esté "confundido", en cada punto se trata de otro espacio.

En ocasiones a los vectores base del espacio tangente inducidos por un sistema local de coordenadas se les denota por

$$(\frac{\partial}{\partial q^1}, \cdots, \frac{\partial}{\partial q^n}). \tag{3.287}$$

Los vectores tangentes pueden verse como operadores, esto es, como aquellos operadores que al aplicarlos a funciones de la variedad en R nos dan la derivada direccional de la función en la dirección indicada por ese vector. Desde ese punto de vista se suele expresar la operación del vector $v_x \in TN_x$ sobre la función $f : N \to R$ por $v_x[f]$. En una carta local, al proyectar por medio del homeomorfismo, la derivada direccional en la dirección de un

vector tangente estará dada por

$$v_x[f] = \sum_{i=1}^{n} v_x(q^i) \left.\frac{\partial f}{\partial q^i}\right|_x , \tag{3.288}$$

donde $v_x(q^i)$ son las componentes del vector tangente v_x en la base inducida por las coordenadas. Esto nos lleva a una nueva notación para la base $(e_1, \cdots, e_n)$, la denotaremos de ahora en adelante por

$$\left(\frac{\partial}{\partial q^1}, \cdots, \frac{\partial}{\partial q^n}\right), \tag{3.289}$$

y a las componentes de un vector tangente v_x en esta base por v_x^i, así

$$v_x[f] = \left(\sum_{i=1}^{n} v_x^i \frac{\partial}{\partial q^i}\right)[f] = \sum_{i=1}^{n} v_x^i \left.\frac{\partial f}{\partial q^i}\right|_x . \tag{3.290}$$

El haber definido los espacios tangentes a las variedades diferenciales abre la posibilidad de construir campos vectoriales sobre variedad, esto es asignar a cada punto de una variedad un vector del espacio tangente en ese punto y además nos permite construir los conceptos de continuidad y diferenciabilidad de los campos vectoriales. Para hacer esto último resulta útil construir una variedad diferencial a partir de todos los espacios tangentes, como tenemos un espacio tangente R^n en cada punto de la variedad, esta nueva variedad es $N \times R^n$ con una estructura diferencial construida a partir de la estructura diferencial de la variedad original N, no desarrollaremos aquí los detalles de la construcción que pueden ser encontrados en los textos especializados, a esta variedad, que tiene dimensión $2n$ se le llama el haz tangente a la variedad y se denota por TN.

Nótese que el haz tangente de la variedad de configuración de un sistema mecánico, representa simultáneamente los puntos de la variedad y la 'velocidad' en esos puntos. En una carta inducida en el haz tangente por un sistema local de coordenadas en la variedad de configuración, las coordenadas de un punto son $(q^1, \cdots, q^n, \dot{q}^1, \cdots, \dot{q}^n)$, en este sentido la función Lagrangiana debe entenderse como una función del haz tangente de la variedad de configuración y, posiblemente, del tiempo en R, esto es

$$L : TM \times R \longrightarrow R. \tag{3.291}$$

Es importante hacer notar que al quedar definido el concepto de campo vectorial sobre una variedad diferencial, podemos extender el concepto de

ecuación diferencial a las variedades diferenciales, ya que, como vimos en el primer capítulo, una ecuación diferencial y un campo vectorial son construcciones matemáticas equivalentes.

Finalmente podemos desarrollar el concepto de diferencial de un mapeo entre dos variedades diferenciales. Tomemos un mapeo diferenciable

$$F : M \longrightarrow N, \tag{3.292}$$

donde M y N son variedades diferenciables de dimensión m y n respectivamente. Tomemos un punto $x \in M$ y su correspondiente $F(x) \in N$. Deseamos encontrar una definición de la diferencial de F valuada en x, que denotaremos por DF_x , que responda a la idea intuitiva de la mejor aproximación lineal a F en el punto x. Para esto DF_x debe ser un mapeo lineal y por tanto debe operar entre espacios vectoriales, esos espacios vectoriales son T_xM y $T_{F(x)}N$, esto es la diferencial de F transforma vectores tangentes a M en el punto x en vectores tangentes a N en el punto $F(x)$

$$DF_x : TM_x \longrightarrow TN_{F(x)}. \tag{3.293}$$

Pasemos a ver como opera la diferencial. Tomemos un vector $v \in TM_x$, v es una clase de equivalencia de curvas tangentes en x, tomemos una curva representativa de esta clase, $c : R \to M$ con $c(0) = x$. Usando el mapeo F generamos una curva en N , $F \circ c : R \to N$, que pasa por $F(x)$, esta curva en N pertenece a una de las clases de equivalencia de curvas de N que pasan por $F(x)$ y, por tanto, define un vector tangente a N en $F(x)$, digamos $v^{'}$; la diferencial de F valuada en x es el mapeo que envía v en v'

$$DF_x : v \longmapsto v'. \tag{3.294}$$

Es fácil ver que este mapeo es lineal y que corresponde a la idea intuitiva, además se puede demostrar que DF_x es independiente de las curvas representativas que utilicemos para definirla. Si ahora introducimos un sistema de coordenadas locales en M que comprenda a x y otro en N que incluya a $F(x)$, entonces podemos ver que, al bajar por medio de los homeomorfismos correspondientes los vectores de M y N a R^m y R^n el mapeo inducido por DF_x, queda establecido por la matriz jacobiana del mapeo de R^m en R^n que se obtiene a partir de F por medio de los homeomorfismos de las cartas.

Esto establece la correspondencia conceptual con el caso común de la diferencial de mapeos entre espacios vectoriales.

Hemos en esta sección hecho un resumen de aquellos aspectos de la teoría de las variedades diferenciales de mayor relevancia para la mecánica clásica.

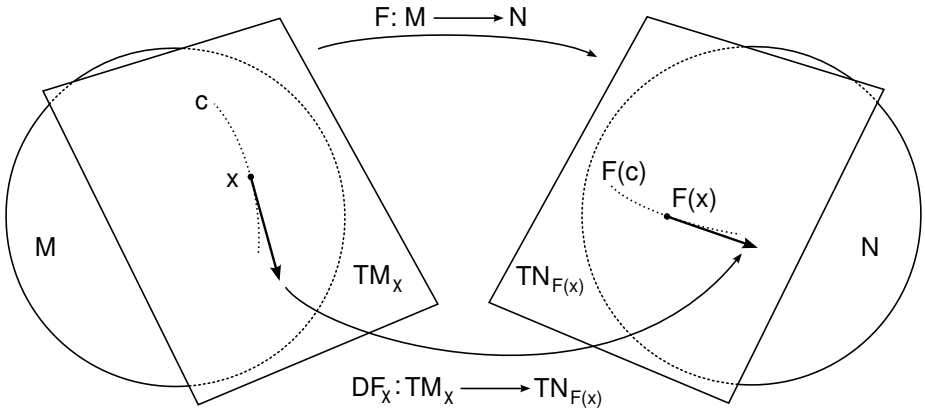


Figura 3.20: La diferencial de un mapeo entre dos variedades diferenciales.

Hemos dejado de lado, por la naturaleza misma de este texto, los detalles, las cuestiones más formales y muchas derivaciones, demostraciones y consecuencias importantes, el lector interesado podrá encontrar estas cuestiones en cualquiera de los muchos textos especializado en la materia. Desde el punto de vista de la mecánica clásica lo más relevante es hacer ver que los sistemas mecánicos "viven" en variedades diferenciales que, en general, no son espacios cartesianos ni euclidianos. Esto debe tomarse siempre en consideración, pues tiene consecuencias importantes. El desarrollo del cálculo diferencial sobre variedades diferenciales nos permite, sin mayores dificultades, trabajar sobre ellas, además nos hace ver que, en muchos casos, podemos trabajar haciendo uso de un solo sistema local de coordenadas sin cometer errores, siempre y cuando éste abarque una amplia región de la variedad de configuración y seamos conscientes de lo que hacemos y de las posibles consecuencias de la topología de la variedad de configuración.

Capítulo 4

Oscilaciones pequeñas.

En muchas ocasiones encontramos sistemas mecánicos en los que la función potencial presenta, en algún punto de la variedad de configuración, un mínimo. Si ponemos al sistema en la configuración correspondiente a ese mínimo y con velocidades generalizadas nulas, el sistema se encontrará en equilibrio estable; quedará inmóvil en esa situación y, ante una pequeña desviación ya sea en la configuración o en las velocidades, el sistema permanecerá en un entorno de dicho equilibrio. Cuando esto sucede y nos encontramos en una vecindad del punto de equilibrio, el sistema se moverá en cada una de las coordenadas generalizadas con un movimiento oscilatorio de amplitud variable pero pequeña, por esto nos referimos a esos movimientos en torno de un punto de equilibrio estable con la denominación de oscilaciones pequeñas.

En el régimen de oscilaciones pequeñas, los sistemas mecánicos que las poseen, pueden aproximarse por un sistema lineal y éste puede, a su vez, ser visto como la superposición de oscilaciones armónicas periódicas que se conocen como modos normales de oscilación. En este capítulo estudiaremos el comportamiento de los sistemas mecánicos que presentan un mínimo en su función potencial y su generalización a sistemas con un número infinito de grados de libertad, lo que nos permitirá introducirnos a la mecánica de los medios continuos y a la teoría de los campos.

4.1. Modos normales de oscilación.

Consideremos un sistema mecánico de n grados de libertad descrito por unas coordenadas generalizadas $(q^1, q^2, \cdots, q^n)$ y una función potencial $V(q^1, q^2, \cdots, q^n)$ que presenta un mínimo; supondremos, por simplicidad, que el mínimo se encuentra en el origen de coordenadas, de no ser así, sin

perdida alguna de generalidad, podemos trasladar el sistema de coordenadas para que su origen coincida con el mínimo.

Las condiciones para que una función de varias variables tenga un mínimo en el origen son:

1.- Que sus primeras derivadas parciales valuadas en el origen sean nulas

$$\left.\frac{\partial V}{\partial q^i}\right|_0 = 0. \tag{4.1}$$

2.- Que su segunda derivada parcial con respecto a q^1 valuada en el origen sea mayor que cero

$$\left.\frac{\partial^2 V}{\partial q^{1^2}}\right|_0 > 0. \tag{4.2}$$

3.- Que los siguientes determinantes formados por las segundas derivadas parciales valuadas en el origen sean mayores que cero.

$$\left|\begin{array}{cc} \frac{\partial^2 V}{\partial q^{1^2}} & \frac{\partial^2 V}{\partial q^1 \partial q^2} \\ \frac{\partial^2 V}{\partial q^2 \partial q^1} & \frac{\partial^2 V}{\partial q^{2^2}} \end{array}\right|_0 > 0\,,\ \cdots,\ \left|\begin{array}{ccc} \frac{\partial^2 V}{\partial q^{1^2}} & \cdots & \frac{\partial^2 V}{\partial q^1 \partial q^n} \\ & \vdots & \\ \frac{\partial^2 V}{\partial q^n \partial q^1} & \cdots & \frac{\partial^2 V}{\partial q^{n^2}} \end{array}\right|_0 > 0\,. \tag{4.3}$$

En una vecindad, suficientemente pequeña del punto de equilibrio, la función potencial podrá aproximarse por los primeros términos de su desarrollo en serie de potencias. El primer término será una constante, que por tratarse de una función potencial no tiene importancia por lo que lo tomaremos como cero. Los términos lineales son cero por tratarse de un mínimo. Finalmente los primeros términos significativos son los de segundo orden, por lo que el potencial podrá aproximarse por

$$V = \frac{1}{2}\sum_{i,j=0}^{n} V_{ij} q^i q^j, \tag{4.4}$$

donde

$$V_{ij} = \left.\frac{\partial^2 V}{\partial q^i \partial q^j}\right|_0. \tag{4.5}$$

Como el potencial presenta un mínimo en el origen, esta forma bilineal simétrica es positiva definida.

La energía cinética será

$$T = \frac{1}{2}\sum_{i,j=0}^{n} \mathcal{I}_{ij}(q)\dot{q}^i \dot{q}^j, \tag{4.6}$$

que es una forma bilineal en las velocidades generalizadas, simétrica y positiva definida. Debido a que las velocidades generalizadas deben ser pequeñas, a que permanecerán pequeñas y a que las coordenadas permanecerán en una vecindad del origen, podemos, para tener el mismo grado de aproximación en la energía cinética que tenemos en el potencial, aproximar los coeficientes $\mathcal{I}_{ij}(q)$ por sus valores en el origen, de esta forma tanto el potencial como la energía cinética quedarán a segundo orden. Entonces definiendo los coeficientes

$$T_{ij} \equiv \mathcal{I}_{ij}(0), \tag{4.7}$$

tendremos que

$$T = \frac{1}{2}\sum_{i,j=0}^{n} T_{ij}\dot{q}^i\dot{q}^j. \tag{4.8}$$

La lagrangiana del sistema para oscilaciones pequeñas, esto es, para el régimen en donde las aproximaciones que estamos haciendo son validas, será

$$L = \frac{1}{2}\sum_{i,j=0}^{n} T_{ij}\dot{q}^i\dot{q}^j - \frac{1}{2}\sum_{i,j=0}^{n} V_{ij}q^iq^j, \tag{4.9}$$

que dará origen a ecuaciones de movimiento lineales. Nótese que aquí estamos denotando por q a un sistema de coordenadas lineales, pues aunque en general para nuestro sistema las coordenadas generalizadas pueden ser curvilíneas, al tomar las aproximaciones en torno al equilibrio estable, estamos linealizando también las coordenadas en el origen.

Ahora bien en el álgebra lineal se demuestra que, si tenemos dos formas bilineales simétricas y al menos una de ellas es positiva definida, entonces existe un sistema de coordenadas en el que las dos formas bilineales quedan expresadas por matrices diagonales y se pueden escoger las coordenadas de tal forma que la matriz correspondiente a la forma positiva definida sea la matriz unidad.

Así tendremos que, para el sistema (4.9), existe una transformación lineal de coordenadas que nos lleva a unas nuevas coordenadas lineales $(Q^1, Q^2, \cdots, Q^n)$ en las cuales la matriz T_{ij} será la matriz unidad y la matriz V_{ij} será diagonal; denotaremos a los n elementos diagonales de esta última matriz por λ_i. Estos coeficientes λ_i serán todos positivos, ya que estamos en un mínimo del potencial.

En estas coordenadas la lagrangiana (4.9) quedará expresada como

$$L = \frac{1}{2}\sum_{i=1}^{n} \dot{Q}^{i^2} - \frac{1}{2}\sum_{i=1}^{n} \lambda_i Q^{i^2}, \qquad i = 1, \cdots, n, \tag{4.10}$$

las ecuaciones de Lagrange serán

$$\ddot{Q}^i + \lambda_i Q^i = 0, \qquad i = 1, \cdots, n, \tag{4.11}$$

o definiendo $\omega_i^2 \equiv \lambda_i$

$$\ddot{Q}^i + \omega_i^2 Q^i = 0, \qquad i = 1, \cdots, n. \tag{4.12}$$

Estas ecuaciones son ecuaciones de osciladores armónicos con frecuencias ω_i y por tanto sus soluciones son de la forma

$$Q^i = A^i \text{sen}(\omega_i t + \theta_i), \qquad i = 1, \cdots, n \tag{4.13}$$

o en términos de funciones complejas

$$Q^i = C^i e^{\omega_i t} + D^i e^{-\omega_i t}, \tag{4.14}$$

donde los coeficientes A^i, θ_i o C^i, D^i quedan determinados por las condiciones iniciales.

En las coordenadas originales q^i las soluciones se obtendrán aplicando la transformación de coordenadas a las ecuaciones (4.13) o (4.14) y se verán como combinaciones lineales de $\text{sen}(\omega_i t + \theta_i)$ o de $\exp(\omega_i t)$ y $\exp(-\omega_i t)$. Naturalmente (4.13) o (4.14) nos hacen ver que el movimiento en general será la composición, en el espacio de n dimensiones, de n osciladores armónicos independientes; esto es una generalización a n dimensiones de las figuras de Lissajous; naturalmente si las frecuencias ω_i son conmensurables, esto es si existe alguna combinación lineal de ellas con coeficientes enteros que sea cero, el movimiento será periódico y la trayectoria una curva cerrada; si son inconmensurables, como ocurre genéricamente, entonces el movimiento será cuasiperiódico y la trayectoria será abierta.

Si las condiciones iniciales son tales que todos los coeficientes A^i en (4.13) son cero excepto uno de ellos, digamos A^k, el movimiento será una oscilación armónica de la coordenada Q^k con frecuencia ω_k y todas las demás coordenadas permanecerán nulas. En las coordenadas originales oscilarán todas, pero todas con la misma frecuencia ω_k y el movimiento será periódico con frecuencia ω_k. A estos movimientos periódicos, que implican una sola de las frecuencias, se les llama modos normales de oscilación. En correspondencia, a las frecuencias ω_i se les llama frecuencias normales o características y a las coordenadas Q^i coordenadas normales o características. Así, siempre que tengamos un sistema mecánico en el que el potencial presente un mínimo, para movimientos suficientemente pequeños en torno al equilibrio estable

correspondiente, habrá n soluciones periódicas, cada una implicando una sola de las frecuencias normales; a estas soluciones periódicas las llamamos modos normales de oscilación. En general el movimiento del sistema, en el régimen de oscilaciones pequeñas, será una combinación lineal de estos modos normales de oscilación y será periódico o cuasiperiódico en dependencia de la conmensurabilidad de las frecuencias normales. Naturalmente, si externamente excitamos al sistema con un forzamiento, el sistema entrará en resonancia cuando la excitación contenga alguna de las frecuencias normales, entonces la amplitud del modo normal de oscilación correspondiente crecerá exponencialmente. Esto ocurrirá, desde luego, hasta que llegue el momento en que la aproximación lineal ya no sea valida y comiencen a dominar los términos no lineales.

Veamos un método para encontrar la transformación de coordenadas que nos lleva a coordenadas normales y a obtener las frecuencias características de un sistema mecánico que presenta un equilibrio estable.

En esencia lo que debemos hacer es diagonalizar simultáneamente las dos matrices, T_{ij} y V_{ij} , ambas simétricas y la primera de ellas tal que genera una forma bilineal positiva definida. Comencemos por la matriz de energía cinética T; como es simétrica podemos encontrar una transformación ortogonal de coordenadas O_1 que la diagonaliza, esto nos lleva a un sistema de coordenadas $(r^1, r^2, \cdots, r^n)$ dado por

$$r = O_1 q \tag{4.15}$$

en el que la matriz de energía cinética queda como

$$O_1 T O_1^{-1} = \begin{pmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & T_n \end{pmatrix}, \tag{4.16}$$

donde las T_i son los eigenvalores de T y la matriz O_1 está compuesta con los eigenvectores de T y es ortogonal. Como la energía cinética es positiva definida todas las T_i son reales positivos y quedará expresada por

$$T = \frac{1}{2} \sum_{i=1}^{n} T_i \dot{r}^{i^2}. \tag{4.17}$$

En este sistema de coordenadas la matriz del potencial quedará como

$$O_1 V O_1^{-1} = \begin{pmatrix} U_{11} & \cdots & U_{1n} \\ \vdots & & \cdots \\ U_{n1} & \cdots & U_{nn} \end{pmatrix} \tag{4.18}$$

y seguirá siendo simétrica, esto es $U_{ij} = U_{ji}$.

Como todas las T_i son reales positivos podemos escalar cada una de las coordenadas r^i y obtener un nuevo sistema de coordenadas R^i por medio de

$$R^i = r^i \sqrt{T_i}; \tag{4.19}$$

en este sistema de coordenadas

$$\dot{r}^i = \frac{\dot{R}^i}{\sqrt{T_i}} \tag{4.20}$$

y por tanto la energía cinética será

$$T = \frac{1}{2} \sum_{i=1}^{n} \dot{R}^{i^2}; \tag{4.21}$$

la matriz del potencial quedará como

$$\left(\frac{U_{ij}}{\sqrt{T_i}\sqrt{T_j}} \right), \tag{4.22}$$

que continua siendo simétrica.

Ahora podemos diagonalizar con una transformación ortogonal de coordenadas O_2 esta ultima matriz y obtener un nuevo sistema de coordenadas $Q = O_2 R$. En este sistema de coordenadas la energía cinética (4.21) se expresará por

$$T = \frac{1}{2} \sum_{i} \dot{Q}_i^2 \tag{4.23}$$

y el potencial se expresará por medio de una matriz diagonal

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \tag{4.24}$$

donde las λ_i son los eigenvalores del la matriz (4.22), que deberán ser reales positivos debido a que esa matriz representa la linealización en torno de un mínimo del potencial[1]. Estas coordenadas Q son las coordenadas normales,

[1] La matriz del potencial genera también una forma bilineal positiva definida, así todo lo hecho a partir de diagonalizar la matriz de energía cinética puede hacerse comenzando por diagonalizar la matriz del potencial con resultados equivalentes. La simplicidad del cálculo nos indicará en cada ocasión la conveniencia de seguir un camino o el otro.

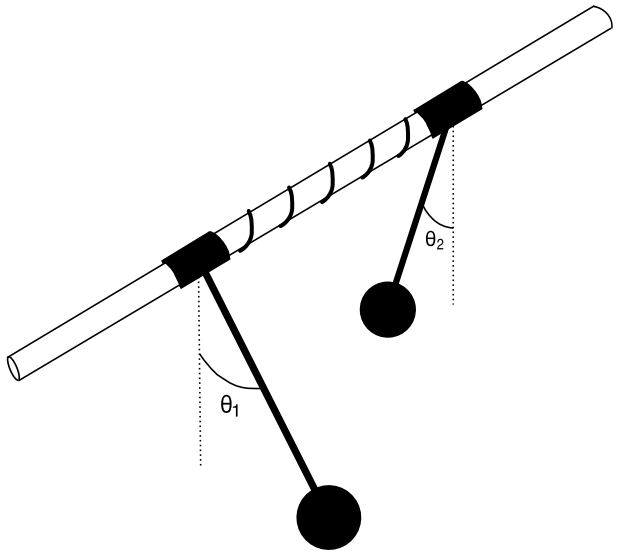


Figura 4.1: Dos péndulos acoplados por un resorte de torsión.

ya que en ellas la matriz de potencial será diagonal y la energía cinética queda expresada por (4.23). Notamos que a estas coordenadas llegamos por medio de la composición de tres transformaciones; primero una transformación ortogonal que diagonaliza la energía cinética; después un escalamiento de cada coordenada que nos lleva a las ecuaciones (4.21) y (4.22); finalmente una transformación ortogonal que diagonaliza la matriz potencial. Notamos que este procedimiento es por si una demostración del teorema al que hicimos referencia.

Veamos un ejemplo. Tomemos dos péndulos simples cuyos puntos de suspensión se encuentran sobre una barra separados por una cierta distancia, consideremos que tienen la misma longitud y masa, l y m. En adición están acoplados por un resorte de torsión que está en reposo cuando los ángulos que subtienden los péndulos con la vertical son iguales, ver figura 4.1.

La energía cinética es

$$T = \frac{1}{2}ml^2\dot{\theta}_1^2 + \frac{1}{2}ml^2\dot{\theta}_2^2 \tag{4.25}$$

y el potencial

$$V = -mgl\cos\theta_1 - mgl\cos\theta_2 + \frac{k}{2}(\theta_1 - \theta_2)^2, \tag{4.26}$$

donde k es la constante torsional del resorte.

Obviamente el potencial tiene un mínimo en $\theta_1 = \theta_2 = 0$, por lo que existe un equilibrio estable cuando ambos péndulos descansan en su posición más baja. Podemos por tanto estudiar la situación para ángulos pequeños y encontrar los modos normales de oscilación. Para hacerlo debemos aproximar el potencial a segundo orden, obtendremos

$$V \approx \frac{(mlg+k)}{2}\theta_1^2 + \frac{(mlg+k)}{2}\theta_2^2 - \frac{k}{2}\theta_1\theta_2 - \frac{k}{2}\theta_2\theta_1 \tag{4.27}$$

y la lagrangiana para oscilaciones pequeñas puesta en la forma (4.9) será

$$L = \frac{1}{2}\left(ml^2\dot{\theta}_1^2 + ml^2\dot{\theta}_2^2\right) - \frac{1}{2}\left((mlg+k)\theta_1^2 + (mlg+k)\theta_2^2 - k\theta_1\theta_2 - k\theta_2\theta_1\right); \tag{4.28}$$

las matrices de energía cinética y potencial serán

$$T = \begin{pmatrix} ml^2 & 0 \\ 0 & ml^2 \end{pmatrix} \tag{4.29}$$

y

$$V = \begin{pmatrix} mlg+k & -k \\ -k & mlg+k \end{pmatrix}. \tag{4.30}$$

La matriz de energía cinética ya es diagonal por lo que no es necesario el primer paso en el proceso de diagonalización simultanea de las dos matrices. Hay que escalar por $\sqrt{ml}$ y las coordenadas que llamamos R serán

$$\begin{aligned} R_1 &= \sqrt{ml}\theta_1, \\ R_2 &= \sqrt{ml}\theta_2, \end{aligned} \tag{4.31}$$

en ellas tendremos que

$$T = \frac{1}{2}(\dot{R}_1^2 + \dot{R}_2^2) \tag{4.32}$$

y la matriz del potencial quedará como

$$U = \begin{pmatrix} \frac{mlg+k}{ml^2} & -\frac{k}{ml^2} \\ -\frac{k}{ml^2} & \frac{mlg+k}{ml^2} \end{pmatrix}. \tag{4.33}$$

Hay que diagonalizar esta última matriz. Nos damos cuenta que es de la forma

$$\begin{pmatrix} a & -b \\ -b & a \end{pmatrix} \tag{4.34}$$

con

$$a = \frac{mlg+k}{ml^2} \qquad y \qquad b = \frac{k}{ml^2}. \tag{4.35}$$

La matriz (4.34) tiene como eigenvalores a $\lambda_1 = a - b$ y $\lambda_2 = a + b$, los eigenvectores normalizados correspondientes son $1/\sqrt{2}(1,1)$ y $1/\sqrt{2}(1,-1)$, la matriz de la transformación de coordenadas que la diagonaliza es

$$O_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{4.36}$$

que es su propia inversa; por tanto las coordenadas normales serán

$$\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \theta_1 + \theta_2 \\ \theta_1 - \theta_2 \end{pmatrix}, \tag{4.37}$$

o a la inversa

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} Q_1 + Q_2 \\ Q_1 - Q_2 \end{pmatrix}. \tag{4.38}$$

Las frecuencias normales serán las raíces de λ_1 y λ_2, las que, una vez hechas las substituciones establecidas en (4.35), resultan ser

$$\omega_1 = \sqrt{\frac{g}{l}} \qquad y \qquad \omega_2 = \sqrt{\frac{g}{l} + \frac{2k}{ml^2}}. \tag{4.39}$$

En coordenadas normales la solución del sistema es

$$Q^1 = A_1 \text{sen}(\omega_1 t + \alpha_1), \tag{4.40}$$

$$Q^2 = A_2 \text{sen}(\omega_2 t + \alpha_2), \tag{4.41}$$

y en las coordenadas originales será

$$\theta_1 = \frac{1}{\sqrt{2}} \left(A_1 \text{sen}(\omega_1 t + \alpha_1) + A_2 \text{sen}(\omega_2 t + \alpha_2) \right), \tag{4.42}$$

$$\theta_2 = \frac{1}{\sqrt{2}} \left(A_1 \text{sen}(\omega_1 t + \alpha_1) - A_2 \text{sen}(\omega_2 t + \alpha_2) \right). \tag{4.43}$$

El primer modo normal de oscilación, el que se obtiene cuando $A_1 \neq 0$ y $A_2 = 0$, es una oscilación simultanea de los dos péndulos en fase con frecuencia ω_1, que es la del péndulo simple, en este caso el resorte no actúa pues los dos ángulos son siempre iguales. El segundo modo normal se da cuando $A_1 = 0$ y $A_2 \neq 0$, vemos que se trata de oscilaciones de frecuencia ω_2 en que los dos péndulos presentan fases contrarias. Notamos que ω_2 es mayor que ω_1 debido a que, en el segundo modo, los péndulos se ven actuados

simultáneamente por la gravedad y por el resorte, esto hace que el sistema sea más rígido que en el primer modo donde únicamente actúa la gravedad.

Este sistema presenta un fenómeno curioso de resonancia cuando el resorte es débil en comparación con la acción de la gravedad, esto es cuando $2k \ll gml$. Consideremos un sistema así en el que, inicialmente, los dos péndulos estén verticales, uno de ellos en reposo y el otro con una velocidad angular inicial Ω, esto es, inicialmente toda la energía del sistema está en uno de los péndulos, digamos el primero. Las condiciones iniciales serán

$$\theta_1 = \theta_2 = \dot{\theta}_2 = 0 \quad y \quad \dot{\theta}_1 = \Omega. \tag{4.44}$$

Esto nos lleva a que las constantes en (4.41) sean

$$\alpha_1 = \alpha_2 = 0, \quad A_1 = \frac{1}{\sqrt{2}}\frac{\Omega}{\omega_1} \quad y \quad A_2 = \frac{1}{\sqrt{2}}\frac{\Omega}{\omega_2} \tag{4.45}$$

por tanto

$$Q_1 = \frac{1}{\sqrt{2}}\frac{\Omega}{\omega_1}\text{sen}\omega_1 t, \tag{4.46}$$

$$Q_2 = \frac{1}{\sqrt{2}}\frac{\Omega}{\omega_2}\text{sen}\omega_2 t. \tag{4.47}$$

Ahora de (4.39) y haciendo la consideración de que k es pequeña podemos poner aproximadamente que

$$\omega_2 = \omega_1 + \frac{k}{\omega_1 ml^2} \equiv \omega_1 + \epsilon, \tag{4.48}$$

donde, por comodidad, hemos llamado ϵ a la cantidad $k/\omega_1 ml^2$. Obtenemos entonces que

$$\theta_1 = \frac{\Omega}{2\omega_1}\left(\text{sen}\omega_1 t + \text{sen}(\omega_1 + \epsilon)t\right), \tag{4.49}$$

$$\theta_2 = \frac{\Omega}{2\omega_1}\left(\text{sen}\omega_1 t - \text{sen}(\omega_1 + \epsilon)t\right). \tag{4.50}$$

Usando ahora las igualdades trigonométricas

$$\text{sen}\varphi = 2\text{sen}\frac{\varphi}{2}\cos\frac{\varphi}{2}, \tag{4.51}$$

$$\cos^2\frac{\varphi}{2} = \frac{1}{2}(\cos\varphi + 1) \tag{4.52}$$

y

$$\text{sen}^2\frac{\varphi}{2} = \frac{1}{2}(1-\cos\varphi), \tag{4.53}$$

llegamos a

$$\theta_1 = \frac{\Omega}{\omega_1}\cos\frac{\epsilon}{2}t\text{sen}(\omega_1+\frac{\epsilon}{2})t, \tag{4.54}$$

$$\theta_2 = -\frac{\Omega}{\omega_1}\text{sen}\frac{\epsilon}{2}t\cos(\omega_1+\frac{\epsilon}{2})t, \tag{4.55}$$

$$\dot\theta_1 = \frac{\Omega\epsilon}{2\omega_1}\cos(\omega_1+\epsilon)t + \Omega\cos\frac{\epsilon}{2}t\cos(\omega_1+\frac{\epsilon}{2})t, \tag{4.56}$$

$$\dot\theta_2 = -\frac{\Omega\epsilon}{2\omega_1}\cos(\omega_1+\epsilon)t + \Omega\text{sen}\frac{\epsilon}{2}t\text{sen}(\omega_1+\frac{\epsilon}{2})t. \tag{4.57}$$

Estas ecuaciones nos indican que los dos péndulos oscilan aproximadamente con la frecuencia $\omega_1+\epsilon/2$, fuera de fase y con una amplitud lentamente variable también fuera de fase. Inicialmente, a $t=0$, tendremos que

$$\theta_1=\theta_2=0, \quad \dot\theta_1 = \Omega + \frac{\Omega\epsilon}{2\omega_1} \quad y \quad \dot\theta_2 = -\frac{\Omega\epsilon}{2\omega_1}, \tag{4.58}$$

que son las condiciones iniciales con una aproximación de orden ϵ y, dentro de esta aproximación, toda la energía está en el péndulo uno. Al tiempo $t=\pi/\epsilon$ tendremos

$$\theta_1 = 0, \quad \theta_2 = \frac{\Omega}{\omega_1}\text{sen}\frac{\omega_1\pi}{\epsilon}, \tag{4.59}$$

$$\dot\theta_1 = -\frac{\Omega\epsilon}{2\omega_1}\text{sen}\frac{\omega_1\pi}{\epsilon} \quad y \quad \dot\theta_2 = -\Omega\cos\frac{\omega_1\pi}{\epsilon} + \frac{\Omega\epsilon}{2\omega_1}\cos\frac{\omega_1\pi}{\epsilon}; \tag{4.60}$$

por lo que, dentro de la aproximación de orden ϵ, la energía se habrá concentrado en el péndulo dos. Desde luego este fenómeno de intercambio de la energía entre los dos péndulos se repetirá periódicamente, esto es, habrá un intercambio periódico de la energía entre los dos péndulos. Esta resonancia no ocurre si los dos péndulos son distintos, ya que las frecuencias serían muy distintas.

4.2. Medios continuos.

Pasemos a analizar otro tipo de sistemas íntimamente ligados a las oscilaciones pequeñas.

Consideremos una cadena lineal de masas ligadas por resortes; por comodidad la tomaremos cerrada, esto es con condiciones periódicas a la frontera,

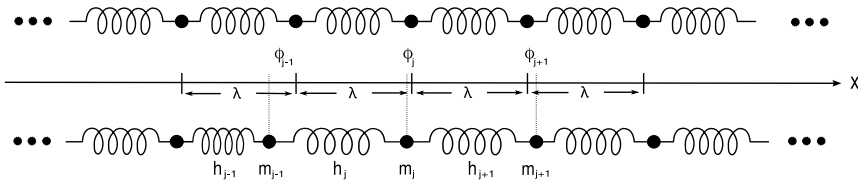


Figura 4.2: Una cadena lineal de masas y resortes. Se utilizan como coordenadas generalizadas las ϕ, esto es, los desplazamientos de cada masa con respecto a la posición de equilibrio representada en la parte superior.

así eliminamos los efectos de las dos masas que quedarían sueltas en los extremos. Denotaremos las masas de las partículas por m_j y las constantes de los resortes por h_j; supondremos que las longitudes en reposo de los resortes son todas iguales a una cantidad λ. De esta forma obtenemos una situación como la ilustrada en la figura 4.2.

Como coordenadas generalizadas utilizaremos el desplazamiento de cada partícula con respecto a su posición de equilibrio cuando el sistema está totalmente relajado y las denotaremos por ϕ_j .

En términos de estas coordenadas las energías cinética y potencial, así como la lagrangiana del sistema se expresan por:

$$T = \frac{1}{2}\sum_j m_j \dot{\phi}_j^2, \tag{4.61}$$

$$V = \sum_j \frac{1}{2} h_j \left(\phi_{j-1} - \phi_j\right)^2 \tag{4.62}$$

y

$$L = \frac{1}{2}\sum_j m_j \dot{\phi}_j^2 - \frac{1}{2}\sum_j h_j \left(\phi_{j-1} - \phi_j\right)^2 . \tag{4.63}$$

Las ecuaciones de movimiento que se obtienen a partir de esta lagrangiana son:

$$m_j \ddot{\phi}_j - h_j \phi_{j-1} + \left(h_{j+1} + h_j\right)\phi_j - h_{j+1}\phi_{j+1} = 0, \tag{4.64}$$

en el caso en que todas las masas y los resortes sean iguales la lagrangiana se reduce a

$$L = \frac{1}{2} m \sum_j \dot{\phi}_j^2 - \frac{1}{2} h \sum_j \left(\phi_{j-1} - \phi_j\right)^2 \tag{4.65}$$

y las ecuaciones de movimiento se reducen a

$$m\ddot{\phi}_j - h\left(\phi_{j-1} - 2\phi_j + \phi_{j+1}\right) = 0. \tag{4.66}$$

En este caso de masas y resortes iguales podemos poner fácilmente la matriz de potencial en la forma (4.22) y obtener

$$\frac{h}{m}\begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & & & & \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 2 \end{pmatrix}, \tag{4.67}$$

donde aparecen dos -1 en los extremos debido a las condiciones de periodicidad en los extremos de la cadena. Al diagonalizar esta matriz obtendríamos las frecuencias y los modos normales.

Por lo pronto no los calcularemos, pues interesa más el estudio de este sistema en el límite continuo cuando el número de partículas tiende a infinito al tiempo que las masas de las partículas y su distancia natural tienden a cero; esto es, cuando $N \to \infty$, $\lambda \to 0$, $m \to 0$ y $h \to \infty$ de tal forma que $N\lambda$ tienda a una longitud finita l, m/λ tienda a una densidad lineal de masa μ y $h\lambda$ a una cantidad finita H. Calculemos la lagrangiana en este límite.

En primer lugar el índice j que enumeraba a las coordenadas generalizadas ϕ_j del sistema discreto pasará a ser un índice continuo que denota la posición de las partículas, ahora infinitesimales, por eso se substituye con una coordenada x que da la posición física a lo largo del sistema, esto es $\phi_j(t)$ pasa a ser $\phi_x(t)$ o $\phi(x,t)$, λ pasa a ser dx y debemos substituir las sumatorias por integrales, por lo que obtendremos para la energía cinética y para el potencial:

$$T = \frac{1}{2}\frac{m}{\lambda}\sum_{j=1}^{N}\lambda\left(\frac{d\phi_j}{dt}\right)^2 \longrightarrow T = \frac{1}{2}\mu\int_0^l\left(\frac{\partial\phi}{\partial t}\right)^2 dx, \tag{4.68}$$

$$V = \frac{1}{2}h\lambda\sum_{j=1}^{n}\lambda\left(\frac{\phi_{j-1}-\phi_j}{\lambda}\right)^2 \longrightarrow V = \frac{1}{2}H\int_0^l\left(\frac{\partial\phi}{\partial x}\right)^2 dx, \tag{4.69}$$

por lo que la lagrangiana será

$$L = \int_0^l\left[\frac{1}{2}\mu\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}H\left(\frac{\partial\phi}{\partial x}\right)^2\right]dx. \tag{4.70}$$

Esta lagrangiana se puede poner como

$$L = \int_0^l \mathcal{L}\left(\phi, \frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial t}\right)dx, \tag{4.71}$$

al definir a la función $\mathcal{L}$ como una densidad lagrangiana, en este caso

$$\mathcal{L} = \frac{1}{2}\mu\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}H\left(\frac{\partial\phi}{\partial x}\right)^2. \tag{4.72}$$

En general cuando tenemos un continuo, ya sea en la mecánica de medios continuos o en la teoría de los campos, tendremos estas densidades lagrangianas y la lagrangiana será la integral sobre el espacio físico de esas densidades lagrangianas. Denotando por (x_1, x_2, x_3) a unas coordenadas cartesianas en el espacio físico tendremos en general que

$$L = \int_V \mathcal{L}\left(\phi, \frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3}, \frac{\partial\phi}{\partial t}\right) d^3r. \tag{4.73}$$

Ahora el principio de Hamilton nos dice que el sistema tendrá el comportamiento que haga que la acción tome un valor extremo, esto es, que la diferencial de la funcional de acción sea cero

$$DS = D\left\{\int_t\int_{x_3}\int_{x_2}\int_{x_1} \mathcal{L}\left(\phi, \frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3}, \frac{\partial\phi}{\partial t}\right) dx_1dx_2dx_3dt\right\} = 0. \tag{4.74}$$

En el capítulo anterior vimos que esto nos conducía a las ecuaciones de Lagrange y el cálculo fue posible gracias a que pudimos introducir el proceso de diferenciación dentro de la integral temporal. Aquí podemos hacer lo mismo pero introduciendo el proceso de diferenciación dentro de la integral temporal y de las integrales espaciales. Siguiendo un proceso totalmente similar al caso discreto obtendremos que para que la diferencial de la acción sea cero se debe cumplir la ecuación

$$\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\frac{\partial\phi}{\partial t}}\right) + \sum_{i=1}^{3}\frac{\partial}{\partial x_i}\left(\frac{\partial\mathcal{L}}{\partial\frac{\partial\phi}{\partial x_i}}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0, \tag{4.75}$$

que es la ecuación de Lagrange para el caso continuo. Naturalmente esta ecuación es una ecuación en derivadas parciales cuya solución, dadas las condiciones iniciales y a la frontera que el sistema físico requiera, nos da ϕ en su variación tanto espacial y como temporal.

Al aplicar esta ecuación a la densidad lagrangiana (4.72) obtenemos

$$\frac{\partial^2\phi}{\partial t^2} - \frac{H}{\mu}\frac{\partial^2\phi}{\partial x^2} = 0, \tag{4.76}$$

que es la ecuación de ondas para una velocidad

$$v = \sqrt{\frac{H}{\mu}}. \tag{4.77}$$

Veamos esto mismo tomando directamente el límite en las ecuaciones de movimiento discretas (4.66). Pongamos estas ecuaciones en la siguiente forma

$$\frac{m}{\lambda}\ddot{\phi}_j - h\lambda\frac{\phi_{j+1}-\phi_j}{\lambda^2} + h\lambda\frac{\phi_j-\phi_{j-1}}{\lambda^2} = 0. \tag{4.78}$$

Al tomar el límite continuo tendremos que

$$\frac{\phi_{j+1}-\phi_j}{\lambda} \to \frac{\phi(x+\lambda)-\phi(x)}{\lambda} \to \left.\frac{\partial\phi}{\partial x}\right|_{x+}, \tag{4.79}$$

$$\frac{\phi_j-\phi_{j-1}}{\lambda} \to \frac{\phi(x)-\phi(x-\lambda)}{\lambda} \to \left.\frac{\partial\phi}{\partial x}\right|_{x-} \tag{4.80}$$

y

$$\frac{\left.\frac{\partial\phi}{\partial x}\right|_{x-} - \left.\frac{\partial\phi}{\partial x}\right|_{x-}}{\lambda} \to \frac{\partial^2\phi}{\partial x^2}, \tag{4.81}$$

por lo que al tomar el tomar el límite la ecuación queda como

$$\mu\frac{\partial^2\phi}{\partial t^2} - H\frac{\partial^2\phi}{\partial x^2} = 0, \tag{4.82}$$

que es precisamente la ecuación (4.76).

Estudiemos los modos normales en este caso continuo. La solución de la ecuación (4.76) debe ser periódica en la longitud del sistema, ya que hemos puesto esas condiciones a la frontera, lo que nos permite escribir que

$$\phi(x,t) = \sum_{n=0}^{\infty}\left[a_n(t)e^{i\frac{2\pi}{l}nx} + a_n^*(t)e^{-i\frac{2\pi}{l}nx}\right], \tag{4.83}$$

o

$$\phi(x,t) = \sum_{n=0}^{\infty}\left[a_n(t)e^{ik_nx} + a_n^*(t)e^{-ik_nx}\right], \tag{4.84}$$

donde el asterisco denota conjugación y hemos definido

$$k_n \equiv \frac{2\pi n}{l}. \tag{4.85}$$

Al sustituir la expresión (4.84) en la ecuación (4.76) encontramos que para que ésta se cumpla se debe cumplir que

$$\ddot{a}_n = -v^2k_n^2a_n \qquad y \qquad \ddot{a}_n^* = -v^2k_n^2a_n^*, \tag{4.86}$$

o

$$\ddot{a}_n = -\omega_n^2 a_n \qquad y \qquad \ddot{a}_n^* = -\omega_n^2 a_n^*, \tag{4.87}$$

donde

$$\omega_n^2 = v^2 k_n^2. \tag{4.88}$$

Por otro lado usando el teorema de Fourier y recordando que la variable x es solo el índice con el que se enumeran las coordenadas generalizadas ϕ_x obtenemos de (4.83)

$$a_n = \frac{1}{l}\int_0^l \phi_x e^{-i(\frac{2\pi}{l})nx} dx = \frac{1}{l}\int_0^l \phi_x e^{-ik_n x} dx, \tag{4.89}$$

esto nos permite ver la relación entre la ϕ_x y las a_n como una transformación de coordenadas generalizadas a coordenadas normales y, por tanto, a las ω_n como las frecuencias normales. Los modos normales serán las soluciones de la ecuación de ondas cuando todas las coordenadas normales a y a^* sean cero excepto unas, digamos las correspondientes al índice n. Entonces las modos normales estarán dados por

$$\phi(x,t) = a_n(t)e^{ik_n x} + a_n^*(t)e^{-ik_n x}, \tag{4.90}$$

o sea por una oscilación armónica de periodo ω_n de las variables ϕ_x, cuyas amplitudes y fases dependen del índice x en la forma dada por (4.90). Como era de esperarse tenemos en este caso un número infinito de modos normales.

En el caso discreto llegamos a los modos normales buscando una transformación de coordenadas generalizadas que nos llevara de la expresión (4.9) para la lagrangiana del sistema a la expresión (4.10) para esta misma lagrangiana. Podemos hacer algo similar en el caso continuo, la lagrangiana en las coordenadas ϕ_x está dada por la ecuación (4.70), calculémosla en las coordenadas normales haciendo uso de las transformaciones (4.89).

La ecuación (4.70) requiere del cálculo de las derivadas parciales de ϕ con respecto a x y a t

$$\frac{\partial\phi}{\partial x} = \frac{\partial}{\partial x}\sum_{n=0}^{\infty}\left[a_n e^{ik_n x} + a_n^* e^{-ik_n x}\right] \tag{4.91}$$

$$= \sum_{n=0}^{\infty} ik_n\left[a_n e^{ik_n x} - a_n^* e^{-ik_n x}\right] \tag{4.92}$$

con respecto a t tendremos

$$\frac{\partial\phi}{\partial t} = \sum_{n=0}^{\infty}\left[\dot{a}_n e^{ik_n x} + \dot{a}_n^* e^{-ik_n x}\right], \tag{4.93}$$

Al calcular los cuadrados de estas derivadas para ser usados en la expresión (4.70) tendremos

$$\left(\frac{\partial\phi}{\partial x}\right)^2 = -\sum_{n,m=0}^{\infty} k_n k_m \Big[a_n a_m e^{i(k_n+k_m)x} + a_n^* a_m^* e^{-i(k_n+k_m)x} - \\ -a_n a_m^* e^{i(k_n-k_m)x} - a_n^* a_m e^{-i(k_n-k_m)x}\Big] \tag{4.94}$$

y

$$\left(\frac{\partial\phi}{\partial t}\right)^2 = \sum_{n,m=0}^{\infty} \Big[\dot{a}_n \dot{a}_m e^{i(k_n+k_m)x} + \dot{a}_n^* \dot{a}_m^* e^{-i(k_n+k_m)x} \\ +\dot{a}_n \dot{a}_m^* e^{i(k_n-k_m)x} + \dot{a}_n^* \dot{a}_m e^{-i(k_n-k_m)x}\Big] \tag{4.95}$$

Ahora debemos hacer las integrales sobre x, pero tenemos las relaciones siguientes

$$\int_0^l e^{i(k_n+k_m)x}dx = \int_0^l e^{i\frac{2\pi}{l}(n+m)x}dx = l\delta_{n,0}\delta_{m,0} \tag{4.96}$$

y

$$\int_0^l e^{i(k_n-k_m)x}dx = \int_0^l e^{i\frac{2\pi}{l}(n-m)x}dx = l\delta_{n,m}, \tag{4.97}$$

por lo que obtendremos que

$$\int_0^l \left(\frac{\partial\phi}{\partial x}\right)^2 dx = 2l\sum_{n=0}^{\infty} k_n^2 a_n a_n^* \tag{4.98}$$

y

$$\int_0^l \left(\frac{\partial\phi}{\partial t}\right)^2 dx = 2l\sum_{n=0}^{\infty} \dot{a}_n \dot{a}_n^*. \tag{4.99}$$

Finalmente de (4.70) tendremos que

$$L = l\sum_{n=0}^{\infty}(\mu\dot{a}_n\dot{a}_n^* - Hk_n^2 a_n a_n^*) = l\mu\sum_{n=0}^{\infty}(\dot{a}_n\dot{a}_n^* - \omega_n^2 a_n a_n^*), \tag{4.100}$$

lagrangiana que nos proporciona las ecuaciones de movimiento en coordenadas normales (4.86).

Un cálculo interesante es el de la energía total de este sistema expresada en coordenadas normales, que sería la suma de las expresiones (4.68) y (4.69) que representan la energía cinética y la potencial utilizando las ecuaciones

(4.98) y (4.99). Pero para hacer esto primero vemos que la solución para las a es

$$a_n = A_n e^{i\omega_n t} + A_n^* e^{-i\omega_n t}, \tag{4.101}$$

por lo que tendremos

$$a_n a_n^* = A_n^2 e^{2i\omega_n t} + A_n^{*2} e^{-2i\omega_n t} + 2A_n A_n^* \tag{4.102}$$

y

$$\dot{a}_n \dot{a}_n^* = -\omega_n^2 \left(A_n^2 e^{2i\omega_n t} + A_n^{*2} e^{-2i\omega_n t} - 2A_n A_n^* \right), \tag{4.103}$$

por lo que al calcular la energía total obtendremos

$$E = 4l\mu \sum_{n=0}^{\infty} \omega_n^2 A_n A_n^*. \tag{4.104}$$

Esto es, la energía total es la suma de las energías en cada modo normal, la que a su vez es proporcional al cuadrado del producto de la frecuencia del modo por su amplitud. En la teoría cuántica las expresiones (4.98) y (4.99) nos proporcionan el operador hamiltaniano en términos de los operadores de creación y aniquilación de los cuantos del campo que corresponden a los modos normales y a la substitución de las coordenadas normales por sus correspondientes operadores. En el caso de la ecuación de ondas elásticas que estamos tratando esto no lleva al concepto de fonón como cuanto del campo y en otros casos al de las diversas partículas que se asocian al campo.

Capítulo 5

Formulación hamiltoniana de la mecánica.

Para determinar totalmente el estado físico de un sistema mecánico de n grados de libertad se requieren $2n$ variables, por lo que resulta sumamente conveniente expresar las ecuaciones de movimiento como un sistema de $2n$ ecuaciones diferenciales de primer orden o como un campo vectorial en $2n$ dimensiones; en contraste con esto las ecuaciones de Lagrange las expresan como ecuaciones de segundo orden en un espacio de n dimensiones, el espacio o variedad de configuración. Naturalmente, cuando tenemos una lagrangiana concreta siempre podremos, después de hacer uso de las ecuaciones de Lagrange y obtener las ecuaciones de movimiento como ecuaciones de segundo orden, ponerlas como una sistema de $2n$ ecuaciones de primer orden, pero en lo general las ecuaciones (****) no quedan expresadas en esta forma y, por tanto, no podemos utilizar muchas de ventajas conceptuales de esta forma de expresión para profundizar el estudio de la mecánica.

En gran medida esto se debe a que utilizamos como variables para especificar el estado del sistema mecánico a q^i y $\dot{q}^i$, esto es, vemos al sistema en TM, el haz tangente a la variedad de configuración. TM tiene $2n$ dimensiones y sus puntos representan completamente el estado del sistema, por lo que podríamos contemplar las ecuaciones en él, sin embargo la estructura de las ecuaciones de la mecánica hace que las ecuaciones de Lagrange no se puedan expresar directamente y en lo general como deseamos.

Ahora bien TM no es la única variedad donde podemos representar los estados de un sistema mecánico; esto es, para determinar totalmente el estado de dicho sistema podemos utilizar las variables q^i y $\dot{q}^i$, pero podemos también utilizar otras. Entre las variables que podemos utilizar están q^i y p_i, donde

las q^i continúan siendo coordenadas locales para la variedad de configuración y las p_i son los impulsos generalizados correspondientes, como dadas las $\dot{q}^i$ podemos encontrar las p_i y viceversa, (q^i, p_i) son suficientes para determinar el estado del sistema.

Así como $(q^i, \dot{q}^i)$ son coordenadas locales para la variedad TM, (q^i, p_i) lo son para otra variedad de $2n$ dimensiones, la que se conoce como el haz cotangente de M o el espacio fase del sistema mecánico. En esta variedad de $2n$ dimensiones si es posible expresar en forma directa y general las ecuaciones de la mecánica como un sistema de $2n$ ecuaciones de primer orden o sea como un campo vectorial.

La formulación hamiltoniana, que desarrollaremos en este capítulo, hace precisamente esto, expresa las ecuaciones de la mecánica como un campo vectorial, o sistema de ecuaciones de primer orden, en el espacio fase.

5.1. Las ecuaciones de Hamilton.

En el capítulo 4 encontramos la cantidad

$$I = \sum_{k=1}^{n} \dot{q}^k \frac{\partial L}{\partial \dot{q}^k} - L, \tag{5.1}$$

que resulta ser una integral de movimiento cuando la lagrangiana no es una función explícita del tiempo. Por otra parte se ha definido el impulso generalizado por

$$p_i = \frac{\partial L}{\partial \dot{q}^i}, \tag{5.2}$$

lo que nos lleva a pensar en formular la mecánica en términos de las variables q y p en vez de las q y $\dot{q}$ utilizando (5.2) para pasar de las variables $\dot{q}$ a las p o viceversa. Tomemos la cantidad 5.1 y expresémosla en términos de q y p

$$I(q,p) = \sum_{k=1}^{n} p_k \dot{q}^k(q,p,t) - L(q, \dot{q}(q,p,t), t), \tag{5.3}$$

donde hemos hecho explicita la dependencia en q y p por medio de la expresión $\dot{q} = \dot{q}(q,p,t)$ obtenida de (5.2).

Calculemos las derivadas parciales de I

$$\frac{\partial I}{\partial p_i} = \dot{q}^i(q,p,t) + \sum_k p_k \frac{\partial \dot{q}^k(p,q,t)}{\partial p_i} - \sum_k \frac{\partial L}{\partial \dot{q}^k} \frac{\partial \dot{q}^k(q,p,t)}{\partial p_i}, \tag{5.4}$$

usando (5.2) obtenemos que

$$\frac{\partial I}{\partial p_i} = \dot{q}^i(q,p,t); \tag{5.5}$$

ahora con respecto a q^i tendremos

$$\frac{\partial I}{\partial q^i} = \sum_k p_k \frac{\partial \dot{q}^k(q,p,t)}{\partial q^i} - \frac{\partial L}{\partial q^i} - \sum_k \frac{\partial L}{\partial \dot{q}^k}\frac{\partial \dot{q}^k(q,p,t)}{\partial q^i}, \tag{5.6}$$

haciendo uso de (5.2) tendremos que

$$\frac{\partial I}{\partial q^i} = -\frac{\partial L}{\partial q^i}; \tag{5.7}$$

similarmente

$$\frac{\partial I}{\partial t} = -\frac{\partial L}{\partial t}. \tag{5.8}$$

Ahora utilizando las ecuaciones de Lagrange en (5.7) tendremos que

$$\frac{\partial I}{\partial q^i} = -\frac{\partial L}{\partial q^i} = -\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = -\dot{p}_i. \tag{5.9}$$

Por tanto si son validas las ecuaciones de Lagrange también lo serán las ecuaciones (5.5) y (5.9) que constituyen una ecuación diferencial en las variables q y p expresada en la forma de un sistema $2n$ de ecuaciones diferenciales de primer orden.

Inversamente si partimos de las variables q y p, de la expresión en esas variables de la cantidad I y de las ecuaciones (5.5) y (5.9) podemos ver, despejando L

$$L(q,\dot{q},t) = \sum_k p_k(q,\dot{q},t)\dot{q}^k - I(q,p(q,\dot{q},t),t), \tag{5.10}$$

y derivando

$$\frac{\partial L}{\partial q^i} = \sum_k \frac{\partial p_k(q,\dot{q},t)}{\partial q^i}\dot{q}^k - \sum_k \frac{\partial I(q,p(q,\dot{q},t),t)}{\partial p_k}\frac{\partial p_k}{\partial q^i} - \frac{\partial I}{\partial q^i} \tag{5.11}$$

$$= \sum_k \frac{\partial p_k}{\partial q^i}\dot{q}^k - \sum_k \dot{q}^k \frac{\partial p_k}{\partial q^i} - \frac{\partial I}{\partial q^i} = -\frac{\partial I}{\partial q^i} = \dot{p}_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \tag{5.12}$$

donde hemos utilizado las ecuaciones (5.5) y (5.9) y recuperado las ecuaciones de Lagrange, o sea si las ecuaciones (5.5) y (5.9) son validas entonces lo son las de Lagrange.

Así hemos demostrado que las ecuaciones diferenciales (5.5) y (5.9) son matemáticamente equivalentes a las de Lagrange, pero están expresadas en las variables q y p y en la forma de un sistema de $2n$ ecuaciones de primer orden. Estas ecuaciones son pues las que estamos buscando y se conocen con el nombre de ecuaciones de Hamilton; así mismo a la cantidad (5.3), cuando se expresa en términos de q y p, la llamaremos función hamiltoniana o simplemente hamiltoniana y la denotaremos con la letra H.

En la formulación lagrangiana caracterizamos el estado de un sistema mecánico por medio de las variables q y $\dot{q}$, esto es con un sistema de coordenadas generalizadas en la variedad de configuración que expresan la situación geométrica del sistema y con las velocidades generalizadas correspondientes; al hacer el cambio de variables a q y p no hemos hecho un simple cambio de coordenadas, las nuevas variables conforman un espacio distinto en el que representar los estados del sistema mecánico. $(q, \dot{q})$ son un sistema de coordenadas locales para el haz tangente de la variedad de configuración, (q, p) no lo son: son un sistema local de coordenadas para otro espacio que llamaremos el espacio fase[1] del sistema mecánico. El haz tangente de la variedad de configuración es también, en el sentido amplio de la teoría de las ecuaciones diferenciales, un espacio fase para las ecuaciones de movimiento, pero en él no podemos expresar en forma general y natural las ecuaciones como un sistema de ecuaciones de primer orden o un campo vectorial y debemos expresarlas como un sistema de ecuaciones de segundo orden, las ecuaciones de Lagrange; claro está, una vez que tenemos las ecuaciones de Lagrange para un caso concreto, siempre podemos expresarlas como un sistema de primer orden en TM. En el espacio fase de la formulación hamiltoniana si lo podemos hacer de manera directa, son las ecuaciones (5.5) y (5.9), que se conocen como ecuaciones de Hamilton, las que, expresadas en la notación habitual, son:

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i}. \end{aligned} \tag{5.13}$$

[1] En el contexto de la teoría de las ecuaciones diferenciales, llamamos espacio fase al espacio en que la ecuación diferencial se expresa como un sistema de ecuaciones de primer orden o aquel espacio en el que las soluciones son curvas parametrizadas por el tiempo que no se intersectan o en el que se puede ver a la ecuación diferencial como un campo vectorial o un flujo. En el contexto de la mecánica clásica se restringe el concepto de espacio fase al espacio generado por las variables q y p, que desde luego, en el sentido de las ecuaciones diferenciales es un espacio fase para las ecuaciones de movimiento, pero en adición, como veremos adelante, tiene otras propiedades importantes.

De esta forma hemos reformulado la mecánica al establecer que un sistema mecánico consta de un espacio fase, el generado por (q, p), una función real de ese espacio fase, y posiblemente del tiempo, que llamamos la función hamiltoniana, y una ecuación diferencial o campo vectorial hamiltoniano construido a partir de la hamiltoniana por medio de (5.13). Hemos encontrado que esta forma nueva de contemplar la mecánica es matemáticamente equivalente a la formulación lagrangiana.

La nueva formulación, a la que llamaremos hamiltoniana, nos permite establecer las ecuaciones de movimiento en la forma de un campo vectorial en el espacio fase y esto, si bien no adiciona nuevos principios o propiedades físicas a las ya presentes en las ecuaciones de Lagrange o en las leyes de Newton, si nos permite extender nuestro entendimiento de la mecánica, profundizar en su estructura conceptual y generar métodos que amplían nuestra capacidad para tratar con los fenómenos mecánicos y, por extensión, con muchos otros fenómenos y teorías en diversas áreas de la física.

Una primera ventaja de expresar las ecuaciones de movimiento como un campo vectorial en el espacio fase, la encontramos en la facilidad con que se pueden calcular las derivadas totales respecto al tiempo de cantidades físicas, esto es, la razón de cambio de esas cantidades a lo largo de la evolución temporal del sistema mecánico. Tomemos una función del espacio fase EF y del tiempo en los reales

$$F : EF \times R \longrightarrow R, \tag{5.14}$$

su derivada con respecto al tiempo será

$$\frac{dF}{dt} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i} \dot{q}^i + \frac{\partial F}{\partial p_i} \dot{p}_i \right) + \frac{\partial F}{\partial t}, \tag{5.15}$$

ahora a lo largo de las trayectorias del sistema las derivadas temporales de las q y las p están dadas por las ecuaciones de Hamilton, entonces

$$\frac{dF}{dt} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q^i} \right) + \frac{\partial F}{\partial t}. \tag{5.16}$$

Esta expresión, nos permite calcular la derivada temporal de cantidades físicas sin tener que resolver las ecuaciones de movimiento y, por tanto, es de gran importancia. Por esto definimos un nuevo concepto, el de paréntesis de Poisson a partir del término sumatorio de (5.16). Tomemos dos funciones del espacio fase, y posiblemente del tiempo, en los reales, digamos F y G, el

paréntesis de Poisson $\{F, G\}$ entre ellas se define por

$$\{F, G\} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q^i} \right), \tag{5.17}$$

notamos que esta definición implica que el paréntesis de Poisson de dos funciones es a su vez una función del espacio fase y el tiempo en los reales, esto es, el paréntesis de Poisson define una operación binaria en el conjunto de todas las funciones. Con la definición de paréntesis de Poisson que hemos dado, la ecuación (5.16) queda como

$$\frac{dF}{dt} = \{F, H\} + \frac{\partial F}{\partial t}. \tag{5.18}$$

Los paréntesis de Poisson tienen una serie de propiedades importantes. En primer lugar son antisimétricos, esto es

$$\{F, G\} = -\{G, F\}, \tag{5.19}$$

lo que, obviamente implica que el paréntesis de Poisson de una función consigo misma es cero. Además son bilineales, esto es, son lineales en sus dos argumentos, propiedad que heredan de la linealidad del operador derivada. A partir de la definición es sencillo demostrar que los paréntesis de Poisson cumplen con la identidad de Jacobi, esto es, si tenemos tres funciones E, F y G, entonces

$$\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} \equiv 0. \tag{5.20}$$

Usando (5.18) vemos que una función del espacio fase en los reales independiente del tiempo, será una integral de movimiento si su paréntesis de Poisson con la hamiltoniana es cero. Ahora aplicando esto a la hamiltoniana misma nos damos cuenta que, por ser el paréntesis de Poisson de una función con sigo misma cero, la hamiltoniana será una integral de movimiento cuando no dependa explícitamente del tiempo. Este resultado es el mismo que obtuvimos en nuestro estudio de la formulación Lagrangiana cuando encontramos que la expresión (5.1) era una integral de movimiento si la lagrangiana no dependía explícitamente del tiempo. Así mismo tendremos que, cuando la energía cinética es una forma cuadrática, entonces la hamiltoniana es la energía total del sistema mecánico y, si es independiente del tiempo, entonces la energía es una integral de movimiento.

La identidad de Jacobi para los paréntesis de Poisson nos permite demostrar que el paréntesis de Poisson de dos integrales de movimiento es a

su vez una integral de movimiento. Tomemos dos integrales de movimiento F y G, entonces $\{F,H\}=0$ y $\{G,H\}=0$, tomando la identidad de Jacobi tendremos que

$$\{H,\{F,G\}\}+\{F,\{G,H\}\}+\{G,\{H,F\}\}=0, \tag{5.21}$$

pero los dos parentesis de la derecha son cero por serlo uno de sus argumentos por lo que

$$\{H,\{F,G\}\}=0 \tag{5.22}$$

y $\{F,G\}$ será una integral de movimiento. Naturalmente esto no nos dice nada sobre la independencia de esta nueva integral con respecto a las dos originales; resulta ser que la mayor parte de las veces la nueva integral resulta ser trivial o dependiente de las otras, esto impide que se genere un número grande de integrales de movimiento independientes a partir de dos de ellas.

Tomemos como ejemplo, para ver como encontrar las ecuaciones de movimiento en el contexto de la formulación hamiltoniana, el del campo central en coordenadas polares. La función lagrangiana está dada por

$$L=\frac{1}{2}m(\dot{r}^2+r^2\dot{\theta}^2)-V(r). \tag{5.23}$$

De aquí utilizando la definición (5.2) obtenemos

$$\begin{aligned} p_r &= m\dot{r}, \\ p_\theta &= mr^2\dot{\theta}, \end{aligned} \tag{5.24}$$

o, en sentido inverso

$$\begin{aligned} \dot{r} &= \frac{p_r}{m}, \\ \dot{\theta} &= \frac{p_\theta}{mr^2}. \end{aligned} \tag{5.25}$$

Utilizando estas ecuaciones y la definición de la hamiltoniana obtenemos que

$$H=\frac{p_r^2}{2m}+\frac{p_\theta^2}{2mr^2}+V(r) \tag{5.26}$$

y las ecuaciones de Hamilton (5.13) nos dan las ecuaciones de movimiento en el espacio fase

$$\begin{aligned} \dot{r} &= \frac{p_r}{m}, \\ \dot{p}_r &= -\frac{\partial V(r)}{\partial r}+\frac{P_\theta^2}{mr^3}, \\ \dot{\theta} &= \frac{p_\theta}{mr^2}, \\ \dot{p}_\theta &= 0. \end{aligned} \tag{5.27}$$

Estas ecuaciones son equivalentes a las obtenidas en la formulación vectorial o newtoniana **** y a las obtenidas en la formulación lagrangiana *****, naturalmente en la formulación hamiltoniana quedan expresadas como una ecuación diferencial o un campo vectorial en el espacio generado por las variables r, p_r, θ y p_θ, esto es el espacio fase.

5.2. Los teoremas de Liuville y Poincaré.

Consideremos un sistema mecánico con su espacio fase y su función hamiltoniana. En general, a un campo vectorial correspondiente a una ecuación diferencial de la forma general de las ecuaciones de Hamilton, lo llamaremos un campo hamiltoniano, y al flujo que genera un flujo hamiltoniano. Este último será un mapeo del espacio fase en si mismo dependiente de un parámetro, el tiempo,

$$g_H^t : EF \longrightarrow EF \tag{5.28}$$

o en coordenadas

$$g_H^t : (q(0), p(0)) \longmapsto (q(t), p(t)), \tag{5.29}$$

donde $q(t)$ y $p(t)$ son la solución de las ecuaciones de Hamilton que a tiempo cero pasa por $(q(0), p(0))$. Recordamos que el flujo hamiltoniano g_H^t es un grupo de un parámetro de difeomorfismos del espacio fase en si mismo y concentra, en un solo objeto matemático, la totalidad de la evolución temporal del sistema mecánico.

Un flujo hamiltoniano tiene una propiedad importantísima que se recoge en el llamado teorema de Liuville: un flujo hamiltoniano preserva el volumen en el espacio fase, ver figura 5.1.

Veamos esto. Tomemos una ecuación diferencial cualquiera $\dot{x} = f(x)$ en un espacio de n dimensiones. El flujo correspondiente a esta ecuación se podrá escribir como

$$g^t(x) = x + f(x)t + 0(t^2). \tag{5.30}$$

Tomemos ahora una región conexa $D(0)$ del espacio fase y denotemos su volumen por $V(0)$. Esta región ira evolucionando con el flujo a medida que pasa el tiempo y en el tiempo t se habrá transformado en la región $D(t)$ con volumen $V(t)$, esto es

$$D(t) = g^t[D(0)]. \tag{5.31}$$

Calculemos el volumen ocupado por la región $D(t)$.

Notamos que el flujo lo podemos ver como una transformación del espacio fase, esto es como el cambio de cada punto con el tiempo, o como una transformación de coordenadas dependiente del tiempo, ambos quedan

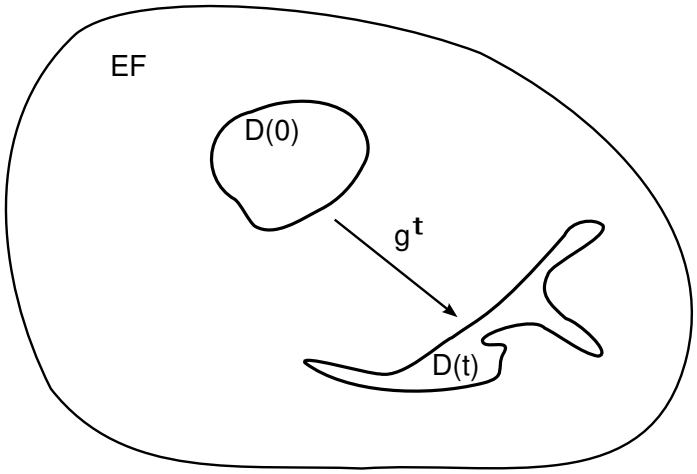


Figura 5.1: La región $D(0)$ se transforma en la región $D(t)$ al aplicar el mapeo resultante de evaluar el flujo al tiempo t. El teorema de Liuville asegura que para un flujo hamiltoniano, si bien la región cambia de forma, el volumen ocupado por ella es invariante.

expresados de la misma manera, por tanto el volumen de la región $D(t)$ puede ser calculado como si se tratara del volumen de la región original $D(0)$ calculado en unas nuevas coordenadas, por tanto

$$V(t) = \int_{D(0)} \det \frac{\partial(g^t(x))}{\partial(x)} d^n x, \tag{5.32}$$

donde

$$\frac{\partial(g^t(x))}{\partial(x)} \tag{5.33}$$

es la matriz jacobiana del flujo al tiempo t. Ahora por ser g^t el flujo de una ecuación diferencial podemos escribir

$$\frac{\partial(g^t(x))}{\partial(x)} = E + \frac{\partial(f(x))}{\partial(x)} t + 0(t^2), \tag{5.34}$$

donde E es la matriz unidad. Ahora, para cualquier matriz A, tenemos que

$$\det(E + At) = 1 + \mathrm{tr}(A)t + 0(t^2), \tag{5.35}$$

donde $\mathrm{tr}(A)$ es la traza de la matriz A, por lo que al aplicar (5.35) a (5.34) encontramos que

$$\det \frac{\partial(g^t(x))}{\partial(x)} = 1 + \mathrm{tr} \frac{\partial(f(x))}{\partial(x)} t + 0(t^2), \tag{5.36}$$

y recordando que

$$\mathrm{tr}\frac{\partial(f(x))}{\partial(x)} = \mathrm{div} f(x), \tag{5.37}$$

obtenemos

$$V(t) = \int_{D(0)} [1 + \mathrm{div} f(x)t] d^n x + 0(t^2), \tag{5.38}$$

por lo que

$$\left.\frac{dV(t)}{dt}\right|_0 = \int_{D(0)} \mathrm{div} f(x) d^n x. \tag{5.39}$$

Vemos entonces que si la divergencia del campo vectorial que define a un flujo es cero, el flujo preservará el volumen del espacio fase.

En el caso de un flujo hamiltoniano el campo vectorial está dado por

$$\left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q}\right) \tag{5.40}$$

y su divergencia es

$$\sum_i \left(\frac{\partial^2 H}{\partial q^i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q^i}\right), \tag{5.41}$$

que vale cero. Por tanto el flujo hamiltoniano preserva el volumen en el espacio fase.

Este resultado y sus consecuencias son de gran importancia, por ejemplo nos hace ver que un sistema hamiltoniano no puede tener atractores, pues de tenerlos una región que convergiera a un atractor tendría necesariamente que disminuir su volumen a medida que evoluciona en el tiempo, cuestión contradictoria con el teorema que acabamos de demostrar. Este resultado cobra importancia capital en la mecánica estadística; en ella implica la conservación de la probabilidad en el "ensamble", por lo que, sin esa propiedad de los flujos hamiltonianos en el espacio fase, la misma formulación básica de mecánica estadística sería inválida.

Es interesante hacer ver que este resultado es consecuencia de las leyes de Newton, ya que la formulación hamiltoniana no introduce nueva física, sólo la formula de manera distinta; pero, desde luego, sería muy enredado expresarlo en términos puramente newtonianos, sin la introducción del espacio fase como sustento de la teoría mecánica.

Estudiemos ahora una consecuencia directa de este teorema, el llamado teorema de recurrencia de Poincaré.

Consideremos un sistema mecánico en el que el espacio fase esté acotado o en el que nos restringimos a una región acotada de él, esto es, una región en la

que las trayectorias permanecen acotadas, por ejemplo el problema de Kepler para energías negativas. Esto implica que ninguna parte del sistema puede llegar a valores infinitos. Tomemos un punto cualquiera de este espacio fase, digamos x y una vecindad cualquiera U de él, entonces podemos afirmar que existe un punto de U (no necesariamente x) tal que la trayectoria que pasa por ese punto, digamos al tiempo cero, pasará nuevamente por la vecindad U después de un tiempo finito T. Pasemos a demostrar esto.

Tomemos el flujo hamiltoniano g_H^t correspondiente a nuestro sistema mecánico y un intervalo de tiempo arbitrario t_0, entonces se genera un mapeo del espacio fase en si mismo $g_H^{t_0}$ que, por simplicidad de la notación, llamaremos g. Este mapeo, por provenir de un flujo hamiltoniano, será diferenciable y preservará el volumen en el espacio fase, por tanto mapeará a la vecindad U en un conjunto $g(U)$ que será abierto, conexo y tendrá el mismo volumen que U. Ahora apliquemos el mapeo a este nuevo conjunto, obtendremos otro conjunto con las mismas propiedades, esto es, será abierto, conexo y tendrá el mismo volumen. Podemos continuar iterando la aplicación del mapeo, obtendremos una sucesión de conjuntos con las propiedades de ser abiertos, conexos y tener todos el mismo volumen

$$U,\ g(U),\ g^{(2)}(U),\ g^{(3)}(U),\ \cdots . \tag{5.42}$$

El espacio fase o la región acotada de este en que estamos interesados, tiene un volumen finito, por lo que deben existir al menos dos de estos conjuntos con intersección no nula, de lo contrario no cabrían en ese volumen finito, ver figura 5.2; por tanto existen dos enteros distintos k y l tales que

$$g^{(k)}(U) \cap g^{(l)}(U) \neq \phi. \tag{5.43}$$

Suponiendo que $l < k$ tendremos que

$$g^{(k-l)}(U) \cap U \neq \phi. \tag{5.44}$$

Tomemos una trayectoria que comience en un punto de U que pertenezca a esta última intersección, después de un tiempo $(k-l)t_0$ estará de nuevo en U, con lo que hemos demostrado la proposición de Poincaré.

Este resultado adquiere gran importancia física cuando estudiamos a un sistema macroscópico como un conjunto de partículas, por ejemplo un vaso totalmente aislado que contiene agua y un hielo. En este caso las partículas son las moléculas de agua y tendremos una cierta hamiltoniana y un espacio fase, como la energía total es finita, ninguna molécula podrá adquirir impulsos infinitos y como el vaso con su tapa ocupa un volumen finito, ninguna de

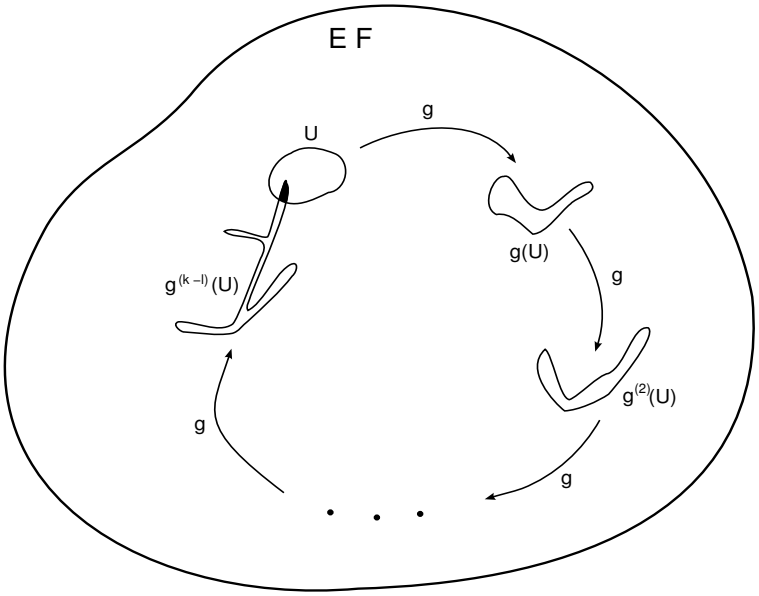


Figura 5.2: Después de aplicar un número finito de veces el mapeo g a una región U, esta se intersecta con la región resultante.

las partículas podrá irse al infinito, por lo que el espacio fase de este sistema está acotado. Dado que la situación inicial es la de tener agua con un hielo, existe una pequeña vecindad del punto del espacio fase que representa esa situación inicial tal que sus puntos representan también estados macroscópicos que veremos como agua con un hielo. El teorema de recurrencia afirma que existe un punto en esta vecindad tal que, al transcurrir el tiempo, su trayectoria se aleja de la vecindad, lo que macroscópicamente significa que el hielo se derrite, pero, después de un tiempo finito, la trayectoria regresa a la vecindad original, esto es, se vuelve a formar el hielo. Claro aquí vemos de inmediato una dificultad con la termodinámica que asegura que el proceso es irreversible ya que la entropía del agua con un hielo es menor que la del agua sin hielo y con temperatura uniforme.

Hay que aclarar que el teorema de recurrencia no dice nada sobre el tiempo que hay que esperar para que se vuelva a formar el hielo, pero se puede estimar. Dado que en este caso el espacio fase tendría del orden de 10^{25} dimensiones, el tiempo de recurrencia sería del orden de 10^{40} años, esto nos hace ver que, para todo fin práctico, observaremos un vaso de agua en equilibrio, pues en la práctica 10^{40} años son muchos años. Por oro lado tendremos que en una formulación de la termodinámica sustentada en la

mecánica clásica, los sistemas acotados y cerrados necesariamente recurren y no sería, en ella, valida la segunda ley. Esto nos hace ver que en todas las formulaciones de la mecánica estadística sustentadas en la mecánica clásica deben introducirse hipótesis adicionales a las leyes de la mecánica para obtener la irreversibilidad.

5.3. EL espacio fase.

Decíamos que el espacio fase y el haz tangente de la variedad de configuración son espacios distintos, para aclarar esto veamos como cambian las velocidades generalizadas $\dot{q}$ al hacer un cambio de las coordenadas q en la variedad de configuración y comparemos con el cambio que ocurre en los impulsos generalizados p al hacer ese mismo cambio de coordenadas.

Tomemos un cambio de coordenadas en la variedad de configuración en el que pasamos de unas coordenadas $(q^1, \ldots, q^n)$ a otras $(Q^1, \ldots, Q^n)$, tendremos entonces una relación biunívoca

$$Q^i = Q^i(q^1, \ldots, q^n), \tag{5.45}$$

$$q^i = q^i(Q^1, \ldots, Q^n); \tag{5.46}$$

de aquí encontramos el cambio que ocurre en las velocidades generalizadas

$$\dot{Q}^i = \sum_k \frac{\partial Q^i}{\partial q^k} \dot{q}^k. \tag{5.47}$$

Pasemos ahora a ver como cambian los impulsos. Por definición $p_i = \partial L/\partial q^i$, así que

$$p_i = \frac{\partial L(Q(q), \dot{Q}(q, \dot{q}))}{\partial \dot{q}^i} \tag{5.48}$$

$$= \sum_k \frac{\partial L}{\partial \dot{Q}^k} \frac{\partial \dot{Q}^k}{\partial \dot{q}^i}, \tag{5.49}$$

pero por definición $P_k = \partial L/\partial \dot{Q}^k$, entonces

$$p_i = \sum_k \frac{\partial \dot{Q}^k}{\partial \dot{q}^i} P_k, \tag{5.50}$$

ahora de (5.47)

$$\frac{\partial \dot{Q}^k}{\partial \dot{q}^i} = \frac{\partial Q^k}{\partial q^i} \tag{5.51}$$

y por tanto

$$p_i = \sum_k \frac{\partial Q^k}{\partial q^i} P_k. \tag{5.52}$$

Comparando (5.47) con (5.52) vemos que las velocidades generalizadas y los impulsos generalizados no se transforman de la misma manera ante una transformación de coordenadas en la variedad de configuración, decimos que unas lo hacen en forma contravariante y los otros en forma covariante; esto es, el cambio de velocidades generalizadas a impulsos generalizados no es un cambio de coordenadas, si lo fuera tendrían las misma propiedades ante transformaciones de coordenadas en la variedad de configuración. Esta es la razón por la que hemos utilizado supraíndices para las velocidades y subíndices para los impulsos e insistido en que el espacio fase no es el haz tangente de la variedad de configuración. Veamos esto con mayor detenimiento[2].

Tomemos un espacio vectorial $\mathbb{V}$ de dimensión n y una base cualquiera para él, $(e_1, e_2, \ldots, e_n)$. Un vector $X \in \mathbb{V}$ se expresará por

$$X = \sum_i x^i e_i, \tag{5.53}$$

donde las x^i son las componentes del vector en la base que hemos escogido.

Consideremos una base distinta a la original $(e_1^{'}, e_2^{'}, \ldots, e_n^{'})$, estará relacionada con la primera por medio de la transformación lineal

$$e_i^{'} = \sum_k \omega_i^k e_k, \tag{5.54}$$

o en forma inversa

$$e_i = \sum_k \rho_i^k e_k^{'}, \tag{5.55}$$

donde naturalmente

$$\sum_k \omega_i^k \rho_k^j = \delta_i^j. \tag{5.56}$$

El vector X se expresará en esta nueva base por

$$X = \sum_i x^{'i} e_i^{'}, \tag{5.57}$$

[2]Con el objeto de entender la estructura y las propiedades del espacio fase como una variedad diferencial simpléctica, iniciamos aquí su estudio y el de la forma en que las leyes naturales, vía las ecuaciones de Hamilton, se relacionan con dicha estructura y propiedades. Esto nos lleva, necesariamente, al estudio de la teoría de las variedades diferenciales, el de los campos vectoriales y el de las formas diferenciales. Para el lector poco acostumbrado al desarrollo matemático puede ser aconsejable hacer una primera lectura, pasar al siguiente tema, y regresar para hacer una segunda lectura más detenida.

donde las $x^{'i}$ son las componentes de X en la nueva base $(e'_1, \ldots, e'_n)$, entonces

$$\sum_i x^i e_i = \sum_k x'^k e'_k = \sum_k x'^k \sum_i \omega_k^i e_i = \sum_i \left[\sum_k x'^k \omega_k^i \right] e_i, \tag{5.58}$$

por lo que tendremos

$$x^i = \sum_k \omega_k^i x^{'k} \tag{5.59}$$

o, inversamente

$$x^{'i} = \sum_k \rho_k^i x^k. \tag{5.60}$$

Nos damos cuenta al comparar (5.54), (5.55), (5.59) y (5.60) que la forma en que se transforman las componentes es inversa a la forma en que lo hacen las bases, por esta razón decimos que, ante un cambio de la base, las componentes contravarían con respecto a la forma en que varía la base, esto lo señalamos utilizando un subíndice para enumerar los vectores de las bases y un supraíndice para numerar las componentes.

Tomemos ahora el conjunto de todas las funciones lineales del espacio vectorial $\mathbb{V}$, las llamadas formas lineales en $\mathbb{V}$, $\varphi : \mathbb{V} \to R$. Si tomamos dos de estas funciones φ y ψ y un escalar, esto es un número real λ, entonces al definir las operaciones producto por un escalar y suma de funciones, $\lambda\varphi$ y $\varphi + \psi$, por medio de

$$(\lambda\varphi)(X) = \lambda(\varphi(X)), \tag{5.61}$$

y

$$(\varphi + \psi)(X) = \varphi(X) + \psi(X), \tag{5.62}$$

el conjunto de las funciones lineales adquiere la estructura de un espacio vectorial; a este espacio se le llama el espacio dual de $\mathbb{V}$ y lo denotaremos por $\mathbb{V}^*$.

Tomemos ahora la base e_i de $\mathbb{V}$, un vector $X \in \mathbb{V}$ y un vector cualquiera $\varphi \in \mathbb{V}^*$, entonces

$$\varphi(X) = \varphi\left(\sum_i x^i e_i \right) = \sum_i x^i \varphi(e_i) = \sum_i x^i a_i, \tag{5.63}$$

donde hemos denotado por a_i a los reales que resultan de aplicar φ a los vectores de la base. Denotemos, ahora, por e^{*i} a n funciones lineales, esto es vectores de $\mathbb{V}^*$, que tengan la propiedad

$$e^{*i}(e_j) = \delta_j^i \tag{5.64}$$

y tomemos la función lineal

$$\sum_i a_i e^{*i}. \tag{5.65}$$

Al aplicarla a X obtenemos

$$\sum_i a_i e^{*i} \left(\sum_j x^j e_j \right) = \sum_{i,j} a_i x^j e^{*i}(e_j) = \sum_{i,j} a_i x^j \delta^i_j = \sum_i a_i x^i, \tag{5.66}$$

el mismo resultado (5.63). Esto nos muestra que los vectores e^{*i} forman una base de $\mathbb{V}^*$, ya que podemos expresar cualquier función lineal de $\mathbb{V}^*$ como una combinación lineal de ellos y, por tanto, el espacio dual de un espacio vectorial es un espacio vectorial de su misma dimensión. A la base e^{*i} se le llama la base dual de e_i. Vemos también que las componentes, en la base dual, de una función lineal o vector de $\mathbb{V}^*$, son los reales que resultan de aplicar esa función a cada uno de los vectores de la base original de $\mathbb{V}$.

Veamos ahora como se comporta la base dual al hacer un cambio de base en el espacio original. Se trata de encontrar la base $e^{'*i}$ dual de la base $e^{'}_i$ dada por (5.54), esto es, $e^{'*i}$ debe ser un conjunto de vectores de $\mathbb{V}^*$ tales que

$$e^{'*i}(e^{'}_j) = \delta^i_j, \tag{5.67}$$

usando (5.54) y (5.56) y el hecho de que debe existir una transformación lineal σ entre las bases e^* y e'^* en $\mathbb{V}^*$ tendremos

$$e^{'*i} = \sum_m \sigma^i_m e^{*m}. \tag{5.68}$$

Ahora, para toda i y toda j,

$$\sum_k \omega^k_j \rho^i_k = \delta^i_j = e^{'*i}(e^{'}_j) = e^{'*i} \sum_k \left(\omega^k_j e_k \right) \tag{5.69}$$

$$= \sum_m \sigma^i_m e^{*m} \sum_k \omega^k_j e_k = \sum_k \omega^k_j \sum_m \sigma^i_m \delta^m_k = \sum_k \omega^k_j \sigma^i_k, \tag{5.70}$$

y por tanto

$$\sigma^i_k = \rho^i_k, \tag{5.71}$$

o sea la base dual se transforma con ρ, esto es, de manera contravariante con respecto a la base original de $\mathbb{V}$. Naturalmente, las componentes de un vector del espacio dual en la base dual se transformarán con ω, esto es en

forma covariante con respecto al cambio de la base original, esto es, si en la base e^{*i} las componentes de una forma lineal ϕ son ϕ_i, en la base e'^{*i} serán

$$\phi'_i = \sum_k \omega_i^k \phi_k. \tag{5.72}$$

Por esta razón hemos utilizado supraíndices para numerar la base dual y subíndices para numerar las componentes de los vectores de $\mathbb{V}^*$ en la base dual. Esto es el sustento básico de la idea de un cambio contravariante y uno covariante, naturalmente con respecto a un cambio de base en el espacio $\mathbb{V}$.

Un espacio vectorial y su dual son espacios vectoriales de la misma dimensión y por tanto son isomorfos, esto es, existen mapeos biunívocos entre ellos que preservan su estructura como espacios vectoriales. De hecho existe una infinidad de dichos mapeos, pero no existe un isomorfismo natural entre ellos. Por otro lado es fácil demostrar que entre un espacio vectorial y el dual de su dual si existe un isomorfismo natural y por tanto podemos identificarlos. Decimos entonces que se trata de un par de espacios duales, ya que en realidad uno es el dual del otro y viceversa. Sin embargo, al no existir un isomorfismo natural entre ellos, no podemos hablar de un vector y su vector dual, sólo podemos referirnos a un espacio vectorial y a su espacio dual y no debemos identificarlos. Ahora cuando en un espacio vectorial se ha definido un producto interno, éste nos permite establecer, independientemente de las bases, un isomorfismo entre él y su dual. Veamos esto con más detenimiento.

Dado un espacio vectorial $\mathbb{V}$ un producto interno es una operación bilineal, simétrica y positiva definida, que asigna a cada par de vectores de dicho espacio un número real; denotaremos esta operación aplicada a los vectores X y Y por $\langle X, Y\rangle$.

Si en el espacio vectorial $\mathbb{V}$ hemos definido una base $(e_1, \ldots, e_n)$ entonces

$$\langle X, Y\rangle = \left\langle \sum_i x^i e_i, \sum_j y^j e_j \right\rangle = \sum_{i,j} \langle e_i, e_j\rangle \, x^i y^j = \sum_{ij} g_{ij} x^i y^j, \tag{5.73}$$

donde hemos hecho uso de la bilinealidad del producto interno y definido $g_{ij} \equiv \langle e_i, e_j\rangle$. Un producto interno nos permite definir una norma en el espacio vectorial por medio de

$$|X|^2 = \langle X, X\rangle = \sum_{ij} g_{ij} x^i x^j \tag{5.74}$$

y por tanto una métrica en el espacio vectorial. Esto nos hace ver que g_{ij} son las componentes de un tensor métrico.

Ahora, como decíamos, el haber definido un producto interno nos permite establecer un isomorfismo entre un espacio vectorial y su dual, para esto asociamos al vector $X \in \mathbb{V}$ el operador o forma lineal $\langle \bullet, X \rangle \in \mathbb{V}^*$ por medio de la definición operativa

$$\langle \bullet, X \rangle [Y] = \langle Y, X \rangle . \tag{5.75}$$

Es fácil ver que esta asociación establece un isomorfismo entre $\mathbb{V}$ y $\mathbb{V}^*$ independiente de las bases, aunque, desde luego, dependiente del producto interno que hayamos definido.

Expresemos al vector $\langle \bullet, X \rangle$ de $\mathbb{V}^*$ en la base dual

$$\langle \bullet, X \rangle = \sum_i a_i e^{*i}, \tag{5.76}$$

donde hemos denotado las componentes por a_i. De las ecuaciones (5.75) y (5.73) tendremos

$$\langle \bullet, X \rangle [Y] = \sum_{ij} g_{ij} x^j y^i, \tag{5.77}$$

pero también tendremos que

$$\langle \bullet, X \rangle [Y] = \sum_i a_i y^i, \tag{5.78}$$

de donde deducimos que

$$a_i = \sum_j g_{ij} x^j, \tag{5.79}$$

o sea el tensor métrico nos sirve para asociar las componentes del vector $\mathbb{V}^*$ en la base dual, a las componentes, en la base original, del vector de $\mathbb{V}$ que tiene asociado por el isomorfismo, por eso se dice, coloquialmente, que el tensor métrico "baja" los índices.

Finalmente señalaremos que si la base que usamos en $\mathbb{V}$ es ortonormal con respecto al producto interno o métrica que hayamos definido, entonces $g_{ij} = \delta_{ij}$ y el isomorfismo se hace transparente, pues resulta que quedan asociados un vector de $\mathbb{V}$ con uno de $\mathbb{V}^*$ cuyas componentes, en la base original unas y en la dual a esta las otras, son los mismos números y por tanto se hacen indistinguibles. Es por esta razón que al usar bases ortonormales se puede confundir un espacio vectorial y su dual y el producto interno se ve como un y producto escalar.

Como habíamos comentado, al comparar las fórmulas de transformación de las componentes de los vectores de un espacio vectorial (5.59) y las de

los vectores de su dual (5.72) con las fórmulas de transformación de las velocidades generalizadas (5.47) y la de los impulsos generalizados (5.50), observamos que hay una enorme semejanza. Las velocidades generalizadas se transforman con la matriz $(\partial Q^i/\partial q^k)$ como si fueran las componentes de un vector en un espacio vectorial, esto es de manera contravariante, mientras los impulsos generalizados lo hacen como si fueran las componentes de los vectores del espacio dual, esto es de manera covariante. Ahora bien, en el caso de la mecánica no tenemos un simple espacio vectorial, pero los conceptos de dualidad desarrollados para espacios vectoriales pueden extenderse a las variedades diferenciales. Para lograr esto, basta hacer en cada espacio tangente a la variedad, lo que hemos hecho en un espacio vectorial, únicamente hay que tener cuidado de que al cambiar de punto, y por tanto de espacio vectorial, las cosas cambien continua y diferencialmente, de lo contrario perderíamos uno de los más importantes conceptos, el de diferenciabilidad. Dada la importancia de todo esto para la mecánica hamiltoniana desarrollaremos a continuación la teoría de los campos vectoriales y las formas diferenciales sobre variedades diferenciales, lo haremos de manera sintética y sin detenernos en muchos detalles o demostraciones, el lector interesado podrá recurrir a los libros especializados en la materia.

En el capítulo tres construimos el haz tangente al unir, en una sola estructura, el conjunto de todos los espacios tangentes a una variedad diferencial y lo hicimos de tal suerte que el haz tangente resultara a su vez una variedad diferencial. Esto nos permitió definir campos vectoriales sobre una variedad diferencial al establecer en cada punto de la variedad un vector del espacio tangente en ese punto y pedir, para que el campo vectorial resulte diferenciable, que, como elementos del haz tangente, al hacer un pequeño cambio en el punto, y tener el correspondiente cambio de vector, el cambio resultase diferenciable. Por otro lado encontramos una forma de asociar a una carta o sistema local de coordenadas $(q^1, \ldots, q^n)$ en la variedad, una base natural para cada espacio tangente. Para ello utilizamos las curvas coordenadas como representantes, en cada punto, de una clase de tangencia y por tanto de un vector del espacio tangente en ese punto, esta base la denotamos por [3]

$$\left(\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}\right). \tag{5.80}$$

En este contexto un campo vectorial se puede expresar dentro de los límites

[3] Es importante recordar que se trata de una base distinta en cada punto ya que cambia el espacio vectorial y además, en general, las curvas coordenadas tienen distintas tangencias en puntos distintos.

de una carta por medio de

$$\sum_i f^i(q^1,\ldots,q^n)\frac{\partial}{\partial q^i}, \tag{5.81}$$

donde las $f^i(q^1,\ldots,q^n)$ son funciones diferenciables de la variedad en los reales a las que llamaremos las componentes del campo vectorial en el sistema de bases asociadas al sistema local de coordenadas. Las $\dot{q}^i$ o velocidades generalizadas denotan, en lo general, las componentes de vectores tangentes en esas bases, pero hay que tener cuidado y no confundirlas con las componentes de un campo vectorial en la variedad de configuración.

Ahora podemos introducir la idea de espacio dual para cada uno de los espacios tangentes a los puntos de una variedad diferencial M. Esto nos lleva a construir para cada punto x de dicha variedad otro espacio vectorial, el dual del espacio tangente en ese punto, cuyos elementos o vectores son formas lineales que aplicadas a un vector tangente (esto es un vector del espacio tangente en el punto) nos dan un número real. Este espacio vectorial lo llamamos el espacio cotangente en el punto y lo denotaremos por TM^*_x. De igual manera que a partir del conjunto de espacios tangentes construimos una nueva variedad diferencial, el haz tangente, a partir del conjunto de espacios cotangentes podemos construir una nueva variedad que llamaremos el haz cotangente y denotaremos por TM^*, naturalmente esta variedad tendrá dimensión doble a la de la variedad inicial, es decir la misma que la del haz tangente.

En el caso de los espacios tangentes, dada una carta de la variedad M y un sistema local de coordenadas $(q^1,\ldots,q^n)$, construimos una base inducida para cada uno de ellos $(\partial/\partial q^1,\ldots,\partial/\partial q^n)$, esto nos permite construir una base inducida para cada uno de los espacios cotangentes tomando la base dual de ésta, denotémosla provisionalmente por

$$\left[\left(\frac{\partial}{\partial q^1}\right)^*,\ldots,\left(\frac{\partial}{\partial q^n}\right)^*\right]. \tag{5.82}$$

Continuando con un desarrollo semejante al que hicimos para el haz tangente, al definir un sistema de coordenadas locales en la variedad de configuración podemos inducir un sistema local de coordenadas en el haz cotangente, para ello tomamos por un lado las coordenadas q del punto de la variedad y, por otro, las componentes de la forma lineal en la base

$$\left[\left(\frac{\partial}{\partial q^1}\right)^*,\ldots,\left(\frac{\partial}{\partial q^n}\right)^*\right], \tag{5.83}$$

o sea $2n$ coordenadas como corresponde al hecho de que el haz cotangente tiene $2n$ dimensiones.

En forma totalmente análoga a la definición de un campo vectorial sobre una variedad diferencial, podemos tomar para cada punto de una variedad diferencial una forma lineal perteneciente al espacio cotangente en el punto y hacerlo de tal manera que el cambio de esta forma lineal al cambiar el punto sea diferenciable dentro de la estructura diferencial del haz cotangente, esto es, construimos un campo de formas lineales sobre la variedad diferencial; a estos campos se les llama uno-formas diferenciales y en unas coordenadas locales q de la variedad se expresan por una fórmula análoga a la (5.81) para los campos vectoriales, esto es por

$$\sum_i f_i(q^1, \ldots, q^n) \left(\frac{\partial}{\partial q^i}\right)^*, \tag{5.84}$$

donde las f_i son funciones diferenciales de la variedad en los reales. Es importante notar que en (5.81) utilizamos supraíndices en las f y aquí utilizamos subíndices, esto es debido a que en el caso de los campos vectoriales las componentes contravarían y en el caso de las uno-formas diferenciales covarían, todo ello con respecto al cambio de la base original

$$\left(\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}\right), \tag{5.85}$$

cambio que está inducido por un cambio de coordenadas en la variedad.

Cuando simultaneamente tenemos una uno-forma diferencial y un campo vectorial definidos en una misma variedad, en cada punto de la variedad tendremos definidos una forma lineal y un vector tangente y podemos aplicar la forma al vector para obtener un número real; podemos decir entonces que las uno-formas diferenciales se aplican u operan sobre los campos vectoriales para dar funciones reales de la variedad en que están definidos.

Dada una variedad diferencial y una función de ella en los reales, una de las uno-formas diferenciales más comunes es la diferencial de la función, ya que, como hemos visto en el capítulo tres, ésta, al evaluarla en un punto de la variedad, es "la mejor aproximación lineal a la función en ese punto", esto es, es un mapeo lineal que se aplica a los vectores tangentes a la variedad en el punto, por tanto, la diferencial de una función es una uno-forma diferencial.

Cuando se establece un sistema local de coordenadas $(q^1, \ldots, q^n)$ en una variedad, cada una de las coordenadas puede verse como una función local de la variedad en los reales, aquella función que dado un punto le asocia el valor de esa coordenada para dicho punto esto es, si tenemos el punto

$x = (q^1, \ldots, q^n)$, entonces $q^i(x) = q^i$. La diferencial de una coordenada será la diferencial de esa función y por tanto será una uno-forma diferencial, veamos cual es el resultado de aplicar esta uno-forma a un vector tangente. Tomemos una de las coordenadas, digamos q^i, su diferencial dq^i y un vector tangente expresado en la base inducida por el sistema local de coordenadas

$$X = \sum_j X^j \frac{\partial}{\partial q^j}. \tag{5.86}$$

Al aplicar la diferencial de una función en un punto a un vector tangente en ese punto, se obtiene la derivada direccional de la función a lo largo del vector, por lo que, si la función es una de las coordenadas, se obtendrá la derivada de la curva coordenada a lo largo del vector, pero $\partial/\partial q^j$ es el vector tangente a la jotaésima curva coordenada, por lo que al aplicar dq^i a X debemos obtener X^i. Esto es

$$dq^i[X] = X^i, \tag{5.87}$$

pero

$$dq^i[X] = dq^i \left[\sum_j X^j \frac{\partial}{\partial q^j}\right] = \sum_j X^j dq^i \left[\frac{\partial}{\partial q^j}\right], \tag{5.88}$$

por lo que debemos tener que

$$dq^i \left[\frac{\partial}{\partial q^j}\right] = \delta^i_j \tag{5.89}$$

y, por tanto, $(dq^1, \ldots, dq^n)$ es, en el espacio cotangente, la base dual de la base $(\partial/\partial q^1, \ldots, \partial/\partial q^n)$ del espacio tangente, esto es, la base que habíamos denotado provisionalmente por

$$\left[\left(\frac{\partial}{\partial q^1}\right)^*, \ldots, \left(\frac{\partial}{\partial q^n}\right)^*\right], \tag{5.90}$$

notación que no volveremos a utilizar.

Así cualquier uno-forma diferencial ϕ se expresa en la base inducida por un sistema local de coordenadas de la variedad por[4]

$$\phi = \sum_j \phi_j(q^1, \ldots, q^n) dq^j. \tag{5.91}$$

[4]Es interesante hacer ver que una uno-forma diferencial es lo que en otra terminología se llama un pffafiano.

En el caso en que se trate de la diferencial de una función F tendremos que

$$\phi_j(q^1,\ldots,q^n) = \frac{\partial F(q^1,\ldots,q^n)}{\partial q^j}. \tag{5.92}$$

En coordenadas locales para la variedad $(q^1,\ldots,q^n)$ y utilizando las bases inducidas en TM_q y TM_q^* una forma diferencial ω y un campo vectorial V se expresarán por

$$\begin{aligned} \omega &= \sum_i \omega_i(q) dq^i, \\ X &= \sum_j X^j(q) \frac{\partial}{\partial q^j}, \end{aligned} \tag{5.93}$$

por lo que el resultado de operar con la forma sobre el campo será

$$\omega[X] = \sum_{ij} \omega_i(q) X^j(q) dq^i \left[\frac{\partial}{\partial q^j}\right] = \sum_i \omega_i(q) X^i(q). \tag{5.94}$$

Por otro lado dado un campo vectorial X sobre un variedad diferencial y una función real F de esa misma variedad, podemos ver al campo vectorial operando sobre la función al definir esta operación como la derivada direccional de la función a lo largo del campo en cada punto de la variedad. Una curva representativa del vector tangente que queda definida por el campo vectorial en cada punto, se expresará en coordenadas por $(q^1(t),\ldots,q^n(t))$; entonces tendremos que la derivada direccional de la función en el punto $x = (q^1(0),\ldots,q^n(0))$ es

$$X[F] = \frac{d}{dt}(F(q^1(t),\ldots,q^n(t))\Big|_0 = \sum_i \frac{\partial F}{\partial q^i}\dot{q}^i\Big|_0, \tag{5.95}$$

pero $\dot{q}^i\big|_0$ es la componente del campo vectorial en el punto q con respecto a la base inducida por el sistema de coordenadas, esto es X^i, por lo que podemos escribir

$$X[F] = \sum_i X^i \frac{\partial F}{\partial q^i} = \left(\sum_i X^i \frac{\partial}{\partial q^i}\right)[F]. \tag{5.96}$$

Esto nos proporciona una forma de calcular las derivadas direccionales o la operación de un campo vectorial sobre una función real y una justificación

para la notación que hemos estando utilizando para los vectores de la base de TM_x asociada a unas coordenadas locales de M.

Ahora, si tenemos una función real F su diferencial es la uno-forma diferencial

$$dF = \sum_i \frac{\partial F}{\partial q^i} dq^i, \tag{5.97}$$

la que opera sobre el campo vectorial X resultando

$$dF[X] = \sum_i \frac{\partial F}{\partial q^i} dq^i \left[\sum_j X^j \frac{\partial}{\partial q^j} \right] = \sum_i \frac{\partial F}{\partial q^i} X^i, \tag{5.98}$$

comparando (5.98) con (5.96) vemos que

$$X[F] = dF[X], \tag{5.99}$$

o sea, es lo mismo operar con un campo vectorial sobre una función, que operar con la diferencial de la función sobre el campo vectorial.

La ecuación (5.75) nos permitió, cuando se tiene definido un producto interno en un espacio vectorial, establecer un isomorfismo natural entre ese espacio y su dual. En el caso de tener una variedad diferencial M y de haber definido un producto interno en TM_q para cada punto $q \in M$ que cambie diferencialmente al cambiar de punto, esta misma ecuación, aplicada ahora en cada punto, nos permite establecer un isomorfismo natural entre el espacio tangente en cada punto de la variedad y el espacio cotangente en ese mismo punto. Esto nos permite establecer una correspondencia, independiente de las coordenadas que se usen, entre campos vectoriales y uno-formas diferenciales, aquella que se reduce, en cada punto, al isomorfismo entre los espacios tangente y cotangente en el punto; desde luego esta relación será dependiente del producto interno que hayamos definido, pero independiente de los sistemas de coordenadas, es decir es una relación global. Veamos esto en coordenadas siguiendo un proceso análogo al que seguimos para el caso simple de los espacios vectoriales.

Tomemos dos campos vectoriales

$$V(q) = \sum_i V^i(q) \left. \frac{\partial}{\partial q^i} \right|_q \qquad y \qquad U(q) = \sum_j U^j(q) \left. \frac{\partial}{\partial q^j} \right|_q \tag{5.100}$$

y un producto interno $< V, U >$ entendido como un producto interno definido en cada espacio tangente, desde luego ahora este producto será una función

de la variedad en los reales. Tendremos entonces que

$$< V, U >= \sum_{ij} V^i(q) U^j(q) < \left.\frac{\partial}{\partial q^i}\right|_q , \left.\frac{\partial}{\partial q^j}\right|_q > \tag{5.101}$$

y denotando al producto interno de los vectores de la base por $g_{ij}(q)$

$$< V, U >= \sum_{ij} g_{ij}(q) V^i(q) U^j(q). \tag{5.102}$$

Ahora relacionamos un campo vectorial V con una uno-forma diferencial asociando a V la uno-forma diferencial $< \bullet, V >$, definida como la uno-forma diferencial que al ser aplicada a un campo vectorial arbitrario W da la función $< W, V >$, es fácil ver que se trata de la uno-forma

$$\omega(q) =< \bullet, V >= \sum_{ij} g_{ij}(q) V^i(q) dq^j, \tag{5.103}$$

esto es, dadas las componentes $V^i(q)$ del campo vectorial en la base inducida por un sistema local de coordenadas, las componentes de la uno-forma diferencial $\omega(q)$ asociada al campo vectorial serán, en la base dual $(dq^1, \ldots, dq^n)$,

$$\omega_j(q) = \sum_i g_{ij}(q) V^i(q). \tag{5.104}$$

Naturalmente esta asociación se puede contemplar en sentido inverso, esto es, el campo vectorial asociado a una uno-forma dada será aquel campo V para el cual $< \bullet, V >$ es esa uno-forma, naturalmente sus componentes estarán dadas por

$$V^i(q) = \sum_j g^{ij}(q) \omega_j(q) \tag{5.105}$$

donde, vistas como matrices, $g_{ij}(q)$ y $g^{ij}(q)$ son inversas. El lector puede ver que $g_{ij}(q)$ es el comúnmente llamado tensor métrico, que en realidad son las componentes de un campo tensorial, ya que está definido como un tensor en el espacio tangente en cada punto de la variedad. Las fórmulas que nos relacionan, por medio de este producto interno, las componentes de un campo vectorial con las de una uno-forma diferencial, son las usuales fórmulas utilizadas para "subir o bajar índices".

A una variedad diferencial sobre la que se ha definido un producto interno en cada uno de sus espacios tangentes, se le llama una variedad rimmaniana, éstas tendrán una métrica, esto es tiene sentido en ellas el concepto de distancia entre dos puntos y su estudio es el objeto de la geometría diferencial. No

abundaremos aquí en este tema; pero continuaremos avanzando en el estudio de las formas diferenciales, por estar esto más directamente relacionado con la mecánica clásica.

Las formas diferenciales se pueden definir en más de un orden, así definimos las dos-formas diferenciales como los campos de operadores bilineales (en cada punto) y antisimétricos, que se aplican a un par de campos vectoriales para obtener una función de la variedad en los reales[5]. En cada punto de la variedad el conjunto de todas las dos-formas, que no son otra cosa que los operadores bilineales y antisimétricos en el espacio tangente al punto, generan un espacio vectorial cuya dimensión, debido a la antisimetría, es $n(n-1)/2$. En este espacio la suma de dos de estas formas ω y α queda definida por

$$(\omega + \alpha)[V, U] = \omega[V, U] + \alpha[V, U] \tag{5.106}$$

donde V y U son vectores tangentes, de forma análoga se define la multiplicación por un escalar. Al hacer esto punto por punto queda definida la suma de dos-formas diferenciales y el producto de una de ellas por un escalar.

Se definen así mismo las k-formas diferenciales como campos de operadores que se aplican a k campos vectoriales para dar una función real de la variedad, con la propiedad de antisimetría ante el intercambio de cualquier par de campos vectoriales y la de linealidad en cada uno de sus argumentos cuando se les analiza punto por punto en los espacios tangentes. Las k-formas diferenciales tienen un papel importante en el estudio de las variedades diferenciales, en particular permiten la definición de las integrales de linea, área o volumen en general. En adición, como veremos, las dos-formas diferenciales tiene un papel relevante en la mecánica hamiltoniana. Utilizaremos letras griegas para denotar las k-formas diferenciales.

Una manera de obtener dos-formas a partir de dos uno-formas es la operación llamada producto exterior, que se define así:

Dadas dos uno-formas ω y α, su producto exterior $\omega \wedge \alpha$ es la dos-forma que aplicada a dos campos vectoriales arbitrarios V y U da

$$\omega \wedge \alpha[V, U] \quad = \quad \det \begin{vmatrix} \omega[V] & \alpha[V] \\ \omega[U] & \alpha[U] \end{vmatrix}. \tag{5.107}$$

Podemos ver fácilmente que, en efecto, el producto exterior de dos uno-formas es una dos-forma. Esta definición se puede extender y definir el producto exterior de dos formas de cualquier orden para obtener una forma de orden

[5]Notamos que el producto interno es un campo de formas bilineales, pero simétricas, mientras que las dos-formas son antisimétricas. Recordamos que en general las formas bilineales son los campos tensoriales.

igual a la suma de los ordenes de las formas originales. Denotando las formas por α y β tendremos que

$$[\alpha \wedge \beta]\,(X_1, X_2, \ldots, X_{n+m}) \tag{5.108}$$

$$= \sum (-1)^{\upsilon} \alpha(X_{i_1}, X_{i_2}, \ldots, X_{i_n}) \beta(X_{j_1}, X_{j_2}, \ldots, X_{j_m}), \tag{5.109}$$

donde $i_1, i_2, \ldots, i_n, j_1, j_2, \ldots, j_m$ es una permutación de los números $1, 2, \ldots, n+m$, la suma se hace bajo las condiciones $i_1 < i_2 < \cdots < i_n$ y $j_1 < j_2 < \cdots < j_m$ y υ es uno si la permutación es par o cero si es impar. Podemos ver fácilmente que la definición (5.107) concuerda con la general dada en (5.109) y que el producto exterior es una operación anticonmutativa, distributiva y asociativa.

La ecuación (5.91) nos permite expresar cualquier uno-forma diferencial en términos de las uno-formas básicas dq^i, de la misma manera cualquier dos-forma diferencial podrá expresarse como

$$\omega = \sum_{i<j} \omega_{ij}(q) dq^i \wedge dq^j, \tag{5.110}$$

y en general una k-forma diferencial por

$$\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1, \ldots, i_k}(q) dq^{i_1} \wedge \cdots \wedge dq^{i_k}. \tag{5.111}$$

Esto se debe a que en cada uno de los espacios cotangentes las diferenciales de las coordenadas locales forman una base para el espacio de las uno-formas, que no es otro que el mismo espacio cotangente; los productos exteriores de dos de dichas diferenciales forman una base para el espacio vectorial de las dos-formas o formas bilineales antisimétricas y los productos exteriores de k diferenciales de las coordenadas, una base para el espacio vectorial de las k-formas. Podemos ver que las funciones $\omega_{i_1, \ldots, i_k}(q)$ que aparecen en (5.111) son lo que comúnmente llamamos un tensor o campo tensorial k veces covariante totalmente antisimétrico.

Otra manera de obtener formas de orden superior a partir de las de orden inferior es la operación de diferenciación exterior. Comencemos por identificar la diferenciación exterior de una función o forma de orden cero con su diferencial, que es una uno-forma, después definimos la diferenciación exterior de una uno-forma diferencial dada por

$$\omega = \sum_i \omega_i(q) dq^i \tag{5.112}$$

por medio de

$$d\omega = \sum_i d\omega_i(q) \wedge dq^i = \sum_{ij} \frac{\partial \omega_i}{\partial q^j} dq^j \wedge dq^i, \tag{5.113}$$

definición que se extiende en forma natural a la diferencial exterior de una k-forma diferencial. Notamos que, debido a la anticonmutatibilidad del producto exterior, podemos escribir

$$d\omega = \sum_{i<j} \left(\frac{\partial \omega_i}{\partial q^j} - \frac{\partial \omega_j}{\partial q^i} \right) dq^j \wedge dq^i. \tag{5.114}$$

Es importante hacer ver que la diferencial exterior de una k-forma que a su vez es la diferencial exterior de una (k-1)-forma es cero, debido esto a que las derivadas parciales con respecto a dos o más coordenadas son invariantes ante el orden en que se toman, mientras que el producto exterior de las diferenciales de las coordenadas es anticonmutativo respecto a ese orden. En general diremos que una k-forma diferencial es cerrada si su diferencial exterior es cero.

Cuando se ha definido en la variedad diferencial un producto exterior en cada uno de sus espacios tangentes podemos, como hemos dicho, establecer un isomorfismo entre los campos vectoriales y las uno-formas diferenciales, esto a su vez, nos permite establecer una correspondencia entre las operaciones del cálculo vectorial, como el gradiente o la divergencia, y la operación general que hemos llamado diferenciación exterior sobre formas diferenciales, veamos esto en el caso singular e importante de tres dimensiones.

Tomemos una variedad diferencial de tres dimensiones en la que se ha definido un producto interno y por tanto un isomorfismo entre cada espacio tangente y su correspondiente espacio cotangente, o sea podemos asociar a cada campo vectorial una uno-forma diferencial. Al tomar un sistema local de coordenadas en la variedad (q^1, q^2, q^3) generamos una base para los espacios tangentes $(\partial/\partial q^1, \partial/\partial q^2, \partial/\partial q^3)$ y su dual para los espacios cotangentes (dq^1, dq^2, dq^3); en estas bases la expresión (5.104) nos relaciona las componentes de un campo vectorial con las de su correspondiente uno-forma diferencial según el isomorfismo que establece el producto interno expresado por el tensor métrico.

Ahora tomemos una función $f(q)$ de una variedad en los reales, si las coordenadas que tomamos son cartesianas, su gradiente será el campo vectorial dado por

$$\mathrm{grad} f(q) = \sum_i \frac{\partial f(q)}{\partial q^i} \frac{\partial}{\partial q^i}, \tag{5.115}$$

y la diferencial de la función $f(q)$ será la uno-forma diferencial

$$df(q) = \sum \frac{\partial f(q)}{\partial q^i} dq^i, \tag{5.116}$$

como en coordenadas cartesianas el tensor métrico es unitario, vemos que el isomorfismo nos correlaciona al gradiente de una función con su diferencial, claro esto lo podemos hacer porque hemos tomado coordenadas cartesianas, cosa que no siempre podemos, de hecho sólo lo podremos hacer si la variedad diferencial es euclidiana, pero esto se puede generalizar y decir que dada cualquier variedad con cualquier métrica, el gradiente de una función de la variedad en los reales es el campo vectorial asociado a la diferencial de la función, según la correspondencia establecida por la métrica. Como la diferencial de una función en cualquier sistema de coordenadas locales siempre esta dada por (5.97) la relación (5.105) nos indica que el gradiente de esa función se expresará por

$$\text{grad} f(q) = \sum_{ij} g^{ij}(q) \frac{\partial f(q)}{\partial q^j} \frac{\partial}{\partial q^i}, \tag{5.117}$$

ecuación que nos proporciona las fórmulas usuales de transformación de las componentes del gradiente en coordenadas curvilíneas, y aquí nos permiten calcular el gradiente de una función sobre una variedad diferencial riemmaniana. Cuando el sistema de coordenadas es ortogonal el tensor métrico es diagonal y se acostumbra denotar sus componentes diagonales por

$$h_i = g_{ii}, \tag{5.118}$$

por lo que

$$g^{ii} = \frac{1}{h_i} \tag{5.119}$$

y la expresión del gradiente se reduce a

$$\text{grad} f(q) = \sum_i \frac{1}{h_i} \frac{\partial f(q)}{\partial q^i} \frac{\partial}{\partial q^i}. \tag{5.120}$$

En el caso de tres dimensiones las dos-formas generan un espacio vectorial de dimensión tres y las tres-formas lo generan de dimensión uno, esto debido a que que en general las k-formas en un espacio de dimensión n generan uno de dimensión $n(n-1)\cdots(n-k)/k!$, por ello en tres dimensiones podemos relacionar el espacio vectorial de las dos-formas en cada espacio tangente con

ese espacio tangente y a las tres formas con las funciones o cero-formas. Así, dado un campo vectorial

$$A = A^1 \frac{\partial}{\partial q^1} + A^2 \frac{\partial}{\partial q^2} + A^3 \frac{\partial}{\partial q^3}, \tag{5.121}$$

la estructura riemmaniana le asocia la uno-forma diferencial

$$\omega_A = \sum_{ij=1}^{3} g_{ij} A^j dq^i, \tag{5.122}$$

que en el caso de coordenadas ortogonales queda como

$$\omega_A = h_1 A^1 dq^1 + h_2 A^2 dq^2 + h_3 A^3 dq^3. \tag{5.123}$$

Ahora en este caso de tres dimensiones también podemos asociar al campo vectorial A una dos forma diferencial α_A; para hacer esto usaremos el triple producto escalar. Tomemos dos campos vectoriales arbitrarios X y Y, la asociación entre A y α es la generada por la ecuación

$$\alpha_A[X, Y] = (A, X, Y), \tag{5.124}$$

donde (A, X, Y) es el triple producto escalar de los tres campos, este producto, naturalmente, se debe entender en cada espacio tangente y recordamos que únicamente se define en espacios de tres dimensiones. Finalmente a una tres-forma diferencial $\beta = f(q) dq^1 \wedge dq^2 \wedge dq^3$, le asociamos la función $f(q)$.

En el caso de un sistema ortogonal de coordenadas la ecuación (5.124) nos proporciona

$$\alpha_A = \sqrt{h_1 h_2 h_3} \left(A^3 dq^1 \wedge dq^2 + A^2 dq^1 \wedge dq^3 + A^1 dq^2 \wedge dq^3 \right), \tag{5.125}$$

mientras que (5.123) nos proporciona la uno-forma diferencial ω_A asociada al campo vectorial A; tomemos la diferencial exterior de esta uno-forma

$$d\omega_A = \left(\frac{\partial h_2 A^2}{\partial q^1} - \frac{\partial h_1 A^1}{\partial q^2} \right) dq^1 \wedge dq^2 \tag{5.126}$$

$$- \left(\frac{\partial h_1 A^1}{\partial q^3} - \frac{\partial h_3 A^3}{\partial q^1} \right) dq^1 \wedge dq^3 \tag{5.127}$$

$$+ \left(\frac{\partial h_3 A^3}{\partial q^2} - \frac{\partial h_2 A^2}{\partial q^3} \right) dq^2 \wedge dq^3, \tag{5.128}$$

vemos que el campo vectorial que correspodería a la dos forma diferencial $d\omega_A$ es

$$\frac{1}{\sqrt{h_1h_2h_3}}\left[\left(\frac{\partial h_3A^3}{\partial q^2}-\frac{\partial h_2A^2}{\partial q^3}\right)\frac{\partial}{\partial q^1}\right. \tag{5.129}$$

$$+\left(\frac{\partial h_1A^1}{\partial q^3}-\frac{\partial h_3A^3}{\partial q^1}\right)\frac{\partial}{\partial q^2} \tag{5.130}$$

$$\left.+\left(\frac{\partial h_2A^2}{\partial q^1}-\frac{\partial h_1A^1}{\partial q^2}\right)\frac{\partial}{\partial q^3}\right], \tag{5.131}$$

que es el rotacional de A.

Tomemos ahora la diferencial de la dos forma que le corresponde a un campo A, esto es la dada por (5.125), tendremos

$$d\alpha_A=\left(\frac{\partial\sqrt{h_1h_2h_3}A_1}{\partial q^1}+\frac{\partial\sqrt{h_1h_2h_3}A_2}{\partial q^2}+\frac{\partial\sqrt{h_1h_2h_3}A_3}{\partial q^3}\right)dq^1\wedge dq^2\wedge dq^3, \tag{5.132}$$

a esta tres-forma le corresponde la función

$$\frac{\partial\sqrt{h_1h_2h_3}A_1}{\partial q^1}+\frac{\partial\sqrt{h_1h_2h_3}A_2}{\partial q^2}+\frac{\partial\sqrt{h_1h_2h_3}A_3}{\partial q^3}, \tag{5.133}$$

esto es, la divergencia del campo A.

Vemos así como podemos utilizar un concepto único, el de diferenciación exterior, para tratar las tres operaciones básicas del cálculo vectorial en tres dimensiones. Esto nos permite expresar cualquier teoría que se formule en términos del cálculo vectorial en tres dimensiones, utilizando formas diferenciales, lo que resulta en gran ventaja, ya que utilizamos únicamente una operación, la diferenciación exterior y tendremos definiciones de las operaciones básicas independientes de coordenadas. [6]

En el espacio fase de la mecánica clásica, que es el haz cotangente de la variedad de configuración, un punto representa totalmente el estado físico de un sistema mecánico. Al utilizar un sistema local de coordenadas

[6]El lector habrá notado que no hemos desarrollado mayormente el cálculo de formas diferenciales, hemos únicamente indicado algunas cuestiones que son de importancia para la mecánica clásica, naturalmente todo esto tiene un fundamento matemático sólido que hemos omitido dado el carácter de este libro. En particular destacamos el hecho de que se puede tratar la teoría de las formas diferenciales sin el uso de coordenadas, lo que asegura su aplicación en variedades diferenciales pues muchas de las propiedades que hemos descrito resultan independientes del sistema local de coordenadas aunque aquí hayamos utilizado coordenadas para estudiarlas. El lector encontrará en la bibliografía referencias a textos que tratan el tema.

$(q^1, \ldots, q^n, p_1, \ldots, p_n)$ un punto del espacio fase puede verse integrado por dos partes, un punto de la variedad de configuración y una uno-forma en el espacio cotangente en ese punto, la primera, asociada al punto $(q^1, \ldots, q^n)$ de la variedad de configuración, representa la situación geométrica del sistema mecánico y la segunda, asociada a la uno-forma con componentes $(p_1, \ldots, p_n)$ la situación cinemática. Esto nos permite construir sobre el espacio fase una uno-forma diferencial de gran importancia esto es una forma lineal en cada espacio cotangente del espacio fase, recordamos que este espacio cotangente tiene dimensión $2n$ donde n es el número de grados de libertad del sistema.[7]

En coordenadas canónicas[8] para el espacio fase, esto es para coordenadas de la forma (q, p), esta uno-forma diferencial se expresa por

$$\omega = \sum_{i=1}^{n} p_i dq^i. \tag{5.134}$$

Observamos en primer lugar que no tiene componentes correspondientes a las uno-formas base dp_i y en segundo lugar que al aplicarla a un vector tangente que únicamente tuviera componentes correspondientes a los vectores base $\frac{\partial}{\partial q^i}$ obtendríamos la función

$$\omega \left[\sum_j \dot{q}^j \frac{\partial}{\partial q^j} \right] = \sum_{ij} p_i \dot{q}^j dq^i \left[\frac{\partial}{\partial q^j} \right] = \sum_i p_i \dot{q}^i \tag{5.135}$$

que es la expresión que utilizamos en (5.3) para definir la hamiltoniana y que, en los casos en que la energía cinética es una función cuadrática, es el doble de esta energía. Recordando la fórmula *** del capítulo *** para el caso en que la energía cinética es una forma cuadrática y tomando al tensor

[7] En rigor, al referirnos en lo que sigue al espacio fase y al uso de las q y las p como coordenadas locales en esa variedad de $2n$ dimensiones, y no a su uso como coordenadas locales de la variedad de configuración las unas y componentes de uno-formas diferenciales las otras, deberíamos usar supraíndices tanto para las q como para las p, pero conservaremos la notación que hemos venido utilizando. Esto nos permite usar supra o subíndices para denotar las componentes de los campos vectoriales en el espacio fase en razón de la asociación de la base a las coordenadas q o p. Naturalmente en lo que sigue, cuando nos referimos al espacio fase, la notación no indica el carácter de covariancia o contravariancia ante los cambios de coordenadas en este espacio de $2n$ dimensiones, sino únicamente ante el cambio de coordenadas en la variedad de configuración.

[8] Más adelante se definirán con mayor amplitud las coordenadas canónicas; por lo pronto todo sistema de coordenadas locales para TM^* en que una mitad de ellas, denotadas por $(q^1, \ldots, q^n)$, sean coordenadas para M y la otra mitad, denotadas por $(p_1, \ldots, p_n)$, sean las componentes, en la base $(dq^1, \ldots, dq^n)$, de las uno-formas diferenciales de M, se considerarán coordenadas canónicas para TM^*.

$\mathcal{I}_{ij}$ como un tensor métrico[9] para la variedad de configuración tendríamos que

$$T = \frac{1}{2}\sum_{ij}\mathcal{I}_{ij}\dot{q}^j\dot{q}^i = \frac{1}{2}\sum_i p_i\dot{q}^i = \frac{1}{2}\sum_{ij}\mathcal{I}^{ij}p_jp_i, \tag{5.136}$$

lo que implica que

$$p_i = \sum_j I_{ij}\dot{q^j} \tag{5.137}$$

y

$$\dot{q}^i = \sum_j I^{ij}p_j. \tag{5.138}$$

Esto es, si usamos al tensor $\mathcal{I}$ como tensor métrico, los impulsos generalizados son las componentes de la uno-forma diferencial que se corresponde con el campo vectorial cuyas componentes son las velocidades generalizadas por medio de la correspondencia que este tensor establece entre campos vectoriales y uno-formas diferenciales.

Ahora a partir de la uno-forma diferencial ω podemos por derivación externa obtener una dos forma que denotaremos por γ, esto es

$$\gamma = d\omega = \sum_{i=1}^{n} dp_i \wedge dq^i. \tag{5.139}$$

Esta dos-forma diferencial es de gran importancia para la mecánica y tiene dos propiedades, por ser la diferencial de una uno-forma diferencial es cerrada, esto es su diferencial es cero, y además dado cualquier campo vectorial distinto de cero, existe otro tal que, al aplicar γ a esa pareja, el resultado es una función distinta de cero, por esto se le llama completa o no degenerada. A una dos forma diferencial con esas dos propiedades se le llama una forma simpléctica, esto es, dado un sistema mecánico siempre podemos definir una forma simpléctica sobre su espacio fase[10], cuando lo hacemos decimos que el espacio fase es una variedad simpléctica.

Así como al definir una métrica sobre una variedad se puede establecer un isomorfismo entre sus espacios tangentes y sus espacios cotangentes, cuando se define sobre una variedad una forma simpléctica ocurre lo mismo. Esto lo hacemos diciendo que la forma simpléctica α induce una correspondencia entre un campo vectorial V con la uno-forma diferencial $\alpha[\bullet, V]$, esto es la uno-forma diferencial que aplicada a un campo vectorial arbitrario X da

[9] Esto lo podemos hacer por ser un tensor simétrico y positivo definido.

[10] Esto se debe a que el espacio fase es el haz cotangente de una variedad diferencial, la de configuración.

la función $\alpha[X,V]$. La bilinealidad nos hace ver que en cada punto de la variedad se ha establecido un isomorfismo entre su espacio tangente y su espacio cotangente[11].

Tomemos ahora un sistema mecánico definido por un espacio fase TM^* con forma simpléctica γ definida por (5.139) y una función hamiltoniana $H : TM^* \to R$, que, en lo que sigue, consideraremos que no depende explicitamente del tiempo. La hamiltoniana es una función que resume las características del sitema y su diferencial puede verse como una generadora de las propiedades diferenciales del sistema; ahora las ecuaciones de movimiento determinan diferencialmente las trayectorias y constituyen un campo vectorial en el espacio fase, esto nos inspira a preguntarnos por el campo vectorial que le corresponde a la diferencial de la hamiltoniana por medio de la forma simpléctica. Aplicaremos γ a un par de campos vectoriales arbitrarios utilizando las coordenadas canónicas (q,p)

$$\begin{aligned}
\gamma[X,V] &= \sum_{j=1}^{n} dp_j \wedge dq^j \left[\sum_{i=1}^{n} \left(X^i \frac{\partial}{\partial q^i} + X_i \frac{\partial}{\partial p_i} \right), \sum_{i=1}^{n} \left(V^i \frac{\partial}{\partial q^i} + V_i \frac{\partial}{\partial p_i} \right) \right] \\
&= \sum_{ij=1}^{n} \det \begin{bmatrix} dp_j \left[\left(X^i \frac{\partial}{\partial q^i} + X_i \frac{\partial}{\partial p_i} \right) \right] & dq^j \left[\left(X^i \frac{\partial}{\partial q^i} + X_i \frac{\partial}{\partial p_i} \right) \right] \\ dp_j \left[\left(V^i \frac{\partial}{\partial q^i} + V_i \frac{\partial}{\partial p_i} \right) \right] & dq^j \left[\left(V^i \frac{\partial}{\partial q^i} + V_i \frac{\partial}{\partial p_i} \right) \right] \end{bmatrix} \\
&= \sum_{i=1}^{n} \det \begin{bmatrix} X_i & X^i \\ V_i & V^i \end{bmatrix} \\
&= \sum_{i=1}^{n} (X_i V^i - X^i V_i).
\end{aligned}$$

Ahora apliquemos la diferencial de la hamiltoniana al mismo campo vectorial X

$$\begin{aligned}
dH[X] &= \sum_{ij=1}^{n} \left(\frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i \right) \left[\left(X^j \frac{\partial}{\partial q^j} + X_j \frac{\partial}{\partial p_j} \right) \right] \quad (5.140) \\
&= \sum_{i=1}^{n} \left(\frac{\partial H}{\partial q^i} X^i + \frac{\partial H}{\partial p_i} X_i \right),
\end{aligned}$$

vemos que el campo V necesario para que el resultado de aplicar la forma simpléctica a X y V sea el mismo que el resultado de aplicar la diferencial

[11] La forma α debe ser no degenerada o completa para evitar que $\alpha[\bullet, V]$ fuese idénticamente cero para algún campo $V \neq 0$, pues de ocurrir esto la relación entre cada espacio tangente con su cotangente no sería un isomorfismo.

de la hamiltoniana al campo X es

$$(V^i, V_i) = \left(\frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i}\right), \tag{5.141}$$

o sea el campo vectorial hamiltoniano, el que visto como ecuación diferencial se corresponde con las ecuaciones de Hamilton. Dada una función hamiltoniana H a este campo vectorial lo denotaremos por V_H.

Hemos encontrado algo de suma importancia, el campo vectorial hamiltoniano, el que genera las trayectorias de un sistema mecánico en el espacio fase, es el correspondiente a la diferencial de la función hamiltoniana conforme al isomorfismo que la forma simpléctica establece entre los espacios tangentes y cotangentes al espacio fase.

Podemos entonces formular la mecánica clásica de la siguiente manera:

Un sistema mecánico hamiltoniano consiste de tres elementos,

1. Un espacio o variedad fase que es el haz cotangente de la variedad de configuración. Un punto de este espacio fase representa totalmente el estado físico del sistema mecánico, esto es, se trata de un espacio de estados para el sistema mecánico.

2. Una forma simpléctica sobre este espacio fase, aquella que se genera naturalmente por ser dicho espacio el haz cotangente de una variedad diferencial.

3. Una función del espacio fase en los reales que llamamos función hamiltoniana. En muchos casos esta función es la energía total del sistema.

La evolución temporal del sistema queda determinada por las trayectorias que en el espacio fase genera el campo vectorial hamiltoniano V_H, que es aquel que se asocia a la diferencial de la hamiltoniana por medio de la ecuación

$$dH = \gamma[\bullet, V_H]. \tag{5.142}$$

Podemos decir que la formulación lagrangiana y la hamiltoniana son duales, pues una se expresa en el haz tangente de la variedad de configuración y la otra en el haz cotangente, al tiempo que la relación entre ellas es la de un isomorfismo entre cada uno de los espacios tangentes y cotangentes o, en un lenguaje distinto, entre espacios de velocidades generalizadas y espacios de impulsos generalizados.

En este contexto podemos definir el paréntesis de Poisson de dos funciones reales del espacio fase F y G sin recurrir a coordenadas como hicimos en (5.17) por medio de

$$\{F, G\} = V_G[F] = dF[V_G], \tag{5.143}$$

ahora bien (**??**), (5.140) y (5.141) nos hacen ver que dF es $\gamma[\bullet, V_F]$, puesto que el resultado de aplicar ambas a un campo vectorial arbitrario es el mismo, por tanto tenemos que

$$\{F, G\} = \gamma[V_G, V_F], \tag{5.144}$$

relación equivalente a las anteriores. Esta última fórmula nos permite comprobar que esta definición coincide con (5.17) cuando utilizamos un sistema de coordenadas canónicas (q, p), ya que en ese caso γ se expresa por (5.139) y entonces

$$\gamma[V_G, V_F] = \sum_i \det \begin{vmatrix} dp_i[V_G] & dq^i[V_G] \\ dp_i[V_F] & dq^i[V_F] \end{vmatrix} = \sum_i \det \begin{vmatrix} -\frac{\partial G}{\partial q^i} & \frac{\partial G}{\partial p_i} \\ -\frac{\partial F}{\partial q^i} & \frac{\partial F}{\partial p_i} \end{vmatrix}, \tag{5.145}$$

que coincide con (5.17). Así podemos decir que el paréntesis de Poisson de F con G es la derivada direccional de F a lo largo del campo vectorial hamiltoniano generado por G o la función resultante de aplicar la diferencial de F al campo vectorial hamiltoniano generado por G o la función que se obtiene al aplicar la forma simpléctica a los campos vectoriales hamiltonianos generados por G y F.

En el caso de un sistema mecánico con hamiltoniana H, la forma (5.144) del paréntesis de Poisson de una función con la hamiltoniana, nos indica que éste es la derivada direccional de la función en la dirección del campo vectorial hamiltoniano, por lo que, si ese paréntesis es cero, lo será dicha derivada direccional y por tanto la función será una integral de movimiento. Aquí hemos deducido esta propiedad de los sistemas hamiltonianos en forma global, esto es para todo el espacio fase sin preocuparnos de los sistemas locales de coordenadas o de su condición de ser canónicos.

Un concepto útil es el del conmutador de dos campos vectoriales. Este es un operador bilineal que asocia a una pareja de campos vectoriales A y B otro campo vectorial que denotamos por $[A, B]$ y llamamos conmutador de A y B, lo definimos por medio del efecto de su aplicación, como operador, a una función arbitraria F

$$[A, B][F] = A[B[F]] - B[A[F]]. \tag{5.146}$$

Es fácil ver que la operación así definida es anticonmutativa y que cumple con la identidad de Jacobi.

Ahora si los dos campos vectoriales son hamiltonianos, digamos X_F y X_G tendremos, utilizando (5.143) y (5.20), que para una función arbitraria

E

$$\begin{aligned}
[X_F, X_G][E] &= X_F[X_G[E]] - X_G[X_F[E]] \\
&= X_F[\{E, G\}] - X_G[\{E, F\}] \\
&= \{\{E, G\}, F\} - \{\{E, F\}, G\} \\
&= \{E, \{G, F\}\} \\
&= X_{\{G,F\}}[E],
\end{aligned}$$

podemos entonces escribir

$$[X_F, X_G] = X_{\{G,F\}}. \tag{5.147}$$

Esto es el conmutador de dos campos vectoriales hamiltonianos es el campo vectorial hamiltoniano generado por el paréntesis de Poisson de las dos funciones generadoras originales tomadas en sentido inverso.

Dado que el paréntesis de Poisson de dos funciones y el conmutador de dos campos vectoriales son operaciones anticonmutativas y cumplen con la identidad de Jacobi, tendremos que tanto el espacio de todas las funciones, como el de todos los campos vectoriales, adquieren, con sus respectivas operaciones, la estructura de un álgebra de Lie.

Pasemos ahora a estudiar una importante relación entre los flujos generados por los campos vectoriales en el espacio fase y la invariancia de la estructura simpléctica.

5.4. El flujo hamiltoniano.

Cuando tenemos un difeomorfismo del espacio fase en si mismo, el espacio y su estructura diferencial quedan invariantes, esto es, tendremos el mismo espacio fase contemplado como variedad diferencial. Ahora el espacio fase tiene, en adición, una estructura simpléctica íntimamente ligada a las propiedades físicas del sistema y ésta puede cambiar ante un difeomorfismo; de ocurrir esto el sistema mecánico cambia y por tanto al plantearnos las ecuaciones de movimiento obtendríamos unas distintas a las obtenidas antes de aplicar el difeomorfismo, estas últimas no serían físicamente válidas. En este sentido para considerar que un sistema físico no cambia ante un difeomorfismo del espacio fase, debemos pedir que no cambie la forma simpléctica.

Veamos con cuidado a que nos referimos al decir que la forma simpléctica no cambie ante un difeomorfismo. Pensemos primero en el caso de una función del espacio fase en los reales y en su cambio ante un difeomorfismo

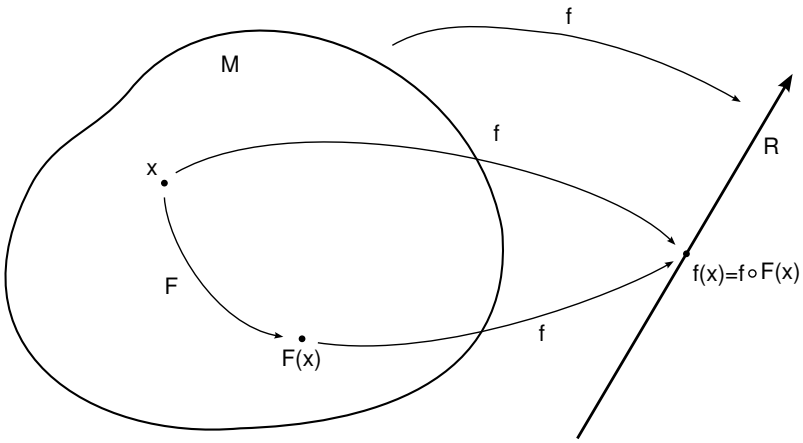


Figura 5.3: Una función f de una variedad diferencial M en los reales es invariante ante un difeomorfismo $F : M \to M$ si $f = f \circ F$.

del espacio. Denotemos por M al espacio fase, por $F : M \to M$ al difeomorfismo y por $f : M \to R$ a la función, como se muestra en la figura 5.3. Podemos ver que al aplicar el difeomorfismo obtendremos una función dada por $f \circ F : M \to R$, o sea, el efecto del difeomorfismo sobre la función f consiste en cambiarla por la función $f \circ F$, que en general es distinta a f. Decimos que la función f es invariante ante el difeomorfismo si ocurre que f valuada en un punto $x \in M$ es igual a f valuada en el punto $F(x)$, esto es, si $f = f \circ F$.

En el caso de una forma simpléctica tenemos que ésta se aplica a pares de campos vectoriales, por lo que antes de definir el efecto de un difeomorfismo sobre una forma simpléctica debemos estudiar su efecto sobre los campos vectoriales.

Tomemos un campo vectorial V sobre M, esto es un vector V_x del espacio tangente TM_x para cada $x \in M$; al aplicar el difeomorfismo obtenemos otro punto de M, el $F(x)$; ahora, recordando la definición de derivada de un mapeo entre variedades diferenciales, podemos asociar al punto $F(x)$ el vector $DF_x[V_x]$, que es un vector del espacio tangente a M en $F(x)$, al hacer esto en cada punto, obtenemos un campo vectorial sobre M; diremos que ese campo vectorial es el resultado de la aplicación del difeomorfismo al campo original y lo denotaremos por $DF[V]$[12].

[12]El lector puede ver que esto es una generalización de lo que comúnmente se hace en R^n, cuando vemos al difeomorfismo como una transformación de coordenadas y a DF como la matriz jacobiana de la transformación.

Ahora una forma diferencial de orden k sobre la variedad M es, en cada punto de la variedad, un operador multilineal antisimétrico que se aplica a k vectores del espacio tangente en ese punto para obtener un número real; podemos decir, al aplicarla punto por punto, que una forma diferencial se puede aplicar a k campo vectoriales y obtener una función de la variedad en los reales, esto es un número en cada punto. En analogía con la transformación de una función de la variedad en los reales ante un difeomorfismo, definiremos el efecto de un difeomorfismo sobre una forma diferencial, diciendo que se obtiene la forma diferencial que, aplicada a k campos vectoriales, da la función que se obtendría al aplicar la forma diferencial original a los k campos vectoriales transformados por el difeomorfismo. Así diremos que el difeomorfismo deja invariante a una forma diferencial si al aplicar el procedimiento anterior obtenemos la misma forma diferencial. En el caso particular que nos interesa, el de la forma simpléctica γ, esto se traduce en la fórmula

$$\gamma[V_1, V_2] = \gamma[DF[V_1], DF[V_2]], \tag{5.148}$$

para todo par de campos vectoriales V_1 y V_2 sobre la variedad, fórmula análoga a

$$f(x) = (f \circ F)(x) = f(F(x)) \tag{5.149}$$

para toda $x \in M$, que establecimos en el caso de las funciones de la variedad en los reales.

Así diremos que el espacio fase es invariante ante un difeomorfismo si y solo si éste deja invariante a la forma simpléctica, esto nos asegura que quedan invariantes las dos componentes del espacio fase: la variedad diferencial y la forma simpléctica.

Diremos que un sistema mecánico hamiltoniano tiene una simetría si existe un grupo de un parámetro $s \in R$ de difeomorfismos del espacio fase en si mismo que deja invariante a la forma simpléctica y a la función hamiltoniana, esto es si para cada valor del parámetro el difeomorfismo que se obtiene deja invariantes a la forma simpléctica y a la función hamiltoniana. En cierto sentido podemos decir que una simetría deja invariante al sistema mecánico, pues deja invariantes todos los elementos que lo definen: el espacio fase, la forma simpléctica y la función hamiltoniana.

Entre la infinidad de difeomorfismos que podemos definir de una variedad en si misma, es importante destacar a los flujos de los campos vectoriales o ecuaciones diferenciales. En la introducción vimos que toda ecuación diferencial tiene como conjunto de sus soluciones un flujo, esto es un grupo de un parámetro de difeomorfismos del espacio fase en si mismo

$$\phi_s : M \to M, \tag{5.150}$$

tales que

$$\phi_{s_1} \circ \phi_{s_2} = \phi_{s_1+s_2}. \tag{5.151}$$

Por otro lado sabemos que todo grupo de un parámetro de difeomorfismos de una variedad en si misma, es el flujo de alguna ecuación diferencial definida en esa variedad, o, lo que es lo mismo, de algún campo vectorial sobre la variedad, campo vectorial que se conoce como el generador infinitesimal del grupo de difeomorfismos. Esto nos hace ver que podemos identificar los conceptos de grupo de un parámetro de difeomorfismos y el de flujo de un campo vectorial.

Cuando un grupo de un parámetro de difeomorfismos o flujo en el espacio fase tiene como generador infinitesimal a un campo vectorial hamiltoniano, se obtienen propiedades interesantes. En especial es de gran importancia el siguiente teorema: Un grupo de un parámetro de difeomorfismos del espacio fase en si mismo (o, en otros términos, un flujo en ese espacio) deja invariante a la forma simpléctica si y sólo si es localmente hamiltoniano.

Pasemos a demostrar este teorema. Tomemos un sistema canónico de coordenadas (q, p) en el espacio fase. Por definición este es un sistema local de coordenadas en el que la forma simpléctica se expresa por

$$\gamma = \sum_i dp_i \wedge dq^i. \tag{5.152}$$

Ahora contemplemos al grupo de difeomorfismos de un parámetro como una transformación de coordenadas a coordenadas (Q^i, P_i), naturalmente esto para cada valor de s. La invariancia de la forma simpléctica se expresa al quedar la forma simpléctica expresada en las nuevas coordenadas por

$$\gamma = \sum_i dP_i \wedge dQ^i, \tag{5.153}$$

para toda s.

El grupo de transformaciones de coordenadas queda expresado por

$$\begin{aligned} q^i &= q^i(Q, P, s) \\ p_i &= p_i(Q, P, s) \end{aligned}$$

o, en sentido inverso, por

$$\begin{aligned} Q^i &= Q^i(q, p, s) \\ P_i &= P_i(q, p, s). \end{aligned}$$

El hecho de que con respecto al parámetro estas transformaciones forman grupo y que para $s = 0$ sean la identidad, nos permite desarrollar en serie

$$\begin{aligned} Q^i &= q^i + \frac{d}{ds}Q^i(q,p,s)\Big|_{s=0}\Delta s + o(\Delta s^2), \\ P_i &= p_i + \frac{d}{ds}P_i(q,p,s)\Big|_{s=0}\Delta s + o(\Delta s^2), \end{aligned}$$

pero las derivadas con respecto al parámetro son las componentes del campo vectorial o, lo que es lo mismo los generadores infinitesimales del grupo, esto es

$$\begin{aligned} Q^i &= q^i + X^i\Delta s + 0(\Delta s^2), \\ P_i &= p_i + X_i\Delta s + 0(\Delta s^2), \end{aligned}$$

donde X^i y X_i son las componentes del campo vectorial en la base inducida por las coordenadas $(\partial/\partial q^i, \partial/\partial p_i)$.

De aquí encontramos que

$$\begin{aligned} dQ^i &= dq^i + \left[\sum_j \left(\frac{\partial X^i}{\partial q^j}dq^j + \frac{\partial X^i}{\partial p_j}dp_j\right)\right]\Delta s + o(\Delta s^2), \\ dP_i &= dp_i + \left[\sum_j \left(\frac{\partial X_i}{\partial q^j}dq^j + \frac{\partial X_i}{\partial p_j}dp_j\right)\right]\Delta s + o(\Delta s^2), \end{aligned}$$

y, por tanto tendremos que

$$\sum_i dP_i \wedge dQ^i = \sum_i dp_i \wedge dq^i \tag{5.154}$$

$$+\Delta s\left[\sum_{ij}\left(\frac{\partial X^i}{\partial q^j} + \frac{\partial X_j}{\partial p_i}\right)dp_i \wedge dq^j + \right. \tag{5.155}$$

$$\left. + \sum_{ij}\left(\frac{\partial X^i}{\partial p_j}dp_i \wedge dp_j + \frac{\partial X_i}{\partial q^j}dq^j \wedge dq^i\right)\right] + o(\Delta s^2). \tag{5.156}$$

Si el campo vectorial es hamiltoniano, esto es si

$$X^i = \frac{\partial H}{\partial p_i} \quad y \quad X_j = -\frac{\partial H}{\partial q^j}, \tag{5.157}$$

tendremos que

$$\sum_{ij}\left(\frac{\partial X^i}{\partial q^j}+\frac{\partial X_j}{\partial p_i}\right)dp_i\wedge dq^j=\sum_{ij}\left(\frac{\partial^2 H}{\partial p_i\partial q^j}-\frac{\partial^2 H}{q^j\partial p_i}\right)dp_i\wedge dq^j=0, \tag{5.158}$$

$$\sum_{ij}\frac{\partial X^i}{\partial p_j}dp_i\wedge dp_j=\sum_{ij}\frac{\partial^2 H}{\partial p_i\partial p_j}dp_i\wedge dp_j=0 \tag{5.159}$$

y

$$\sum_{ij}\frac{\partial X_i}{\partial q^j}dq^j\wedge dq^i=\sum_{ij}\frac{\partial^2 H}{\partial q^j\partial q^i}dq^j\wedge dq^i=0, \tag{5.160}$$

por lo que infinitesimalmente

$$\sum_i dP_i\wedge dQ^i=\sum_i dp_i\wedge dq^i, \tag{5.161}$$

y como las transformaciones forman grupo, la transformación deja invariante a la forma simpléctica.

En sentido inverso si la transformación deja invariante a la forma simpléctica los términos en Δs en (5.156) deben ser cero y por tanto

$$\frac{\partial X^i}{\partial q^j}+\frac{\partial X_j}{\partial p_i}=0\quad\forall ij, \tag{5.162}$$

$$\sum_{ij}\frac{\partial X^i}{\partial p_j}dp_i\wedge dp_j=\sum_{i<j}(\frac{\partial X^i}{\partial p_j}-\frac{\partial X^j}{\partial p_i})dp_i\wedge dp_j=0, \tag{5.163}$$

por lo que

$$\frac{\partial X^i}{\partial p_j}-\frac{\partial X^j}{\partial p_i}=0\qquad para\qquad i<j, \tag{5.164}$$

de igual forma tendremos que

$$\frac{\partial X_i}{\partial q^j}-\frac{\partial X_j}{\partial q^i}=0\qquad para\qquad i<j. \tag{5.165}$$

Ahora la uno-forma diferencial que le corresponde al campo vectorial según la forma simpléctica γ es

$$\gamma[\bullet,X]=\sum_i\left(-X_i dq^i+X^i dp_i\right) \tag{5.166}$$

y su diferencial exterior será

$$\begin{aligned} d\gamma[\bullet, X] &= \sum_i \left[-\sum_j \left(\frac{\partial X_i}{\partial q^j} dq^j + \frac{\partial X_i}{\partial p_j} dp_j \right) \wedge dq^i + \right. \\ &\quad \left. + \sum_j \left(\frac{\partial X^i}{\partial q^j} dq^j + \frac{\partial X^i}{\partial p_j} dp_j \right) \wedge dp_i \right. \\ &= -\sum_{i<j} \left(\frac{\partial X_i}{\partial q^j} - \frac{\partial X_j}{\partial q^i} \right) dq^j \wedge dq^i + \sum_{i<j} \left(\frac{\partial X^i}{\partial p_j} - \frac{\partial X^j}{\partial p_i} \right) dp_j \wedge dp_i + \\ &+ \sum_{ij} \left(\frac{\partial X^i}{\partial q^j} + \frac{\partial X_j}{\partial p_i} \right) dq^i \wedge dp_i, \end{aligned}$$

que es cero dadas las relaciones anteriores. Esto nos lleva a que la uno forma $\gamma[\bullet, X]$ es exacta, por lo que debe ser la diferencial de alguna función, digamos G

$$\gamma[\bullet, X] = dG, \tag{5.167}$$

lo que nos dice que el campo X es hamiltoniano con función hamiltoniana G y queda demostrado el teorema.

Esta demostración es valida únicamente en forma local, puesto que la hemos hecho en un sistema local de coordenadas, para extenderla deberíamos demostrar que en cualquier otro sistema de coordenadas locales se obtiene lo mismo, incluyendo el que, en el segundo sentido, se obtiene la misma función hamiltoniana. La extensión a una demostración global se puede hacer fácilmente en el primer sentido, pero no es posible hacerlo en el segundo cuando el espacio fase no es simplemente conexo. Esto es, el hecho de que un flujo hamiltoniano deja invariante la forma simpléctica es valido globalmente, pero que la invariancia de la forma simpléctica implique que el flujo sea hamiltoniano es válido únicamente de manera local. Esto último es así ya que la función hamiltoniana que genera al flujo queda definida localmente, esto es, únicamente en la carta que estemos usando, y no necesariamente es la misma que generaría dicho flujo en otra carta[13].

[13] Esto no es preocupante ya que siempre podemos tener una función hamiltoniana global en el levantamiento del espacio fase cuando este no sea simplemente conexo.

5.5. El teorema de Noether en la formulación hamiltoniana.

En la versión hamiltoniana del teorema de Noether las transformaciones o difeomorfismos ante las que se considera que un sistema mecánico queda invariante son transformaciones del espacio fase; éstas son mucho más amplias que las transformaciones de la variedad de configuración que consideramos en el contexto de la formulación lagrangiana. En el caso hamiltoniano, a cambio de esa amplitud, debemos agregar algo para poder decir que el sistema mecánico queda invariante ante una transformación del espacio fase, en este caso, en adición a la función hamiltoniana tenemos una forma simpléctica, por lo que debemos pedir que ésta también sea invariante ante la transformación. El teorema de Noether quedará enunciado de la siguiente manera: Si un grupo de un parámetro de difeomorfismos del espacio fase deja invariantes a la forma simpléctica y a la función hamiltoniana, entonces existe una constante de movimiento.

La demostración, una vez demostrado el teorema de la sección anterior, es sencilla y muy clara. Como el grupo deja invariante a la forma simpléctica, su generador infinitesimal es un campo vectorial hamiltoniano, llamemos F a la función hamiltoniana correspondiente. Como el grupo deja invariante a la función hamiltoniana H del sistema mecánico, esta H será una integral de movimiento del grupo, entendido éste como un flujo, por tanto el paréntesis de Poisson de H con F será cero, $\{H, F\} = 0$, pero entonces $\{F, H\} = 0$ y F es una constante de movimiento del sistema mecánico.

Vemos que la constante de movimiento que aparece en el teorema de Noether es la función hamiltoniana del flujo generado por el grupo en el espacio fase y que el teorema nos habla de la correspondencia entre dos flujos hamiltonianos cuando cada uno deja invariante a la función hamiltoniana del otro.

Esto se puede extender al caso general de las acciones de los grupos de Lie sobre el espacio fase. Veamos esto con más cuidado.

Un grupo de Lie es una variedad diferencial G sobre la que se ha definido una operación binaria $* : G \times G \to G$ que forma grupo y que vista como mapeo de $G \times G$ en G es diferenciable. Dados $g, h \in G$ denotaremos a la operación del grupo por $g * h \in G$ y por e al elemento idéntico del grupo, esto es, para toda $g \in G$, $e * g = g * e = g$ y existe $g^{-1} \in G$ tal que $g^{-1} * g = g * g^{-1} = e$. A la dimensión de la variedad diferencial G se le llama la dimensión del grupo de Lie.

Dado un elemento g de un grupo de Lie G y un elemento o punto arbi-

trario $h \in G$, se define el mapeo $L_g : G \to G$ por medio de

$$L_g(h) = g * h, \tag{5.168}$$

este mapeo se conoce como una traslación por la izquierda; podemos también definir las traslaciones por la derecha por medio de

$$R_g(h) = h * g, \tag{5.169}$$

naturalmente, si el grupo es abeliano, ambas traslaciones coinciden; en este texto utilizaremos únicamente las traslaciones por la izquierda.

L_g tiene las siguientes propiedades:

L_g es un mapeo diferenciable.

L_e es el mapeo idéntico.

$L_{g^{-1}} = (L_g)^{-1}$.

Dados $g_1, g_2 \in G$ $L_{g_1} \circ L_{g_2} = L_{g_1 * g_2}$.

Esto implica que L_g define una acción del grupo G sobre la variedad G, o sea, sobre si mismo.

Tomemos el espacio tangente a la variedad G, o sea al grupo, en el punto e, T_eG, y un vector $V \in T_eG$, la diferencial del mapeo L_g valuada en e al aplicarla al vector V nos dará un vector de T_gG, esto es

$$D_eL_g : T_eG \to T_gG \tag{5.170}$$

$$D_eL_g[V] = W \in T_gG; \tag{5.171}$$

por tanto al hacer recorrer g por todo el grupo, D_eL_g genera, para cada vector $V \in T_eG$, un campo vectorial sobre la variedad G con la propiedad que el vector V es el vector del campo en el elemento identidad del grupo, denotaremos a este campo por X_V.

Diremos que un campo vectorial X en G es invariante por la izquierda si

$$D_hL_g[X(h)] = X(g * h) \tag{5.172}$$

para toda $g, h \in G$, esto es, si es invariante ante todos los difeomorfismos L_g. El conjunto de todos los campos vectoriales de G invariantes por la izquierda forma un espacio vectorial ante la suma gracias a la linealidad de D_hL_g y como el conmutador de dos de estos campos será también un campo vectorial invariante por la izquierda, el conjunto adquiere la estructura de un álgebra de Lie.

Ahora dado un vector $V \in T_eG$, X_V es un campo invariante por la izquierda ya que

$$D_hL_g[X_V(h)] = (D_hL_g \circ D_eL_h)[V] \tag{5.173}$$

$$= D_e(L_h \circ L_g)[V] \tag{5.174}$$

$$= D_e L_{h*g}[V] = X_V(h * g). \tag{5.175}$$

Por otro lado dado un campo vectorial X en G invariante por la izquierda le podemos asociar un vector V de T_eG por medio de $V = X(e)$ y se puede ver que el campo vectorial generado por V es X. Así se establece un isomorfismo de espacios vectoriales entre el espacio tangente T_eG y el espacio de todos los campos vectoriales sobre G invariantes por la izquierda; como este último es un álgebra de Lie tomando al conmutador de campo vectoriales como operación, podemos darle a T_eG la estructura de un álgebra de Lie definiendo la operación $\{V, W\}$ entre dos vectores $V, W \in T_eG$ por medio de

$$\{V, W\} = [X_V, X_W](e), \tag{5.176}$$

lo que implica que

$$[X_V, X_W] = X_{\{V,W\}}; \tag{5.177}$$

al espacio T_eG con la operación $\{.,.\}$ así definida se le conoce como el álgebra de Lie del grupo G, se le denota por $\mathfrak{g}$ y es isomorfa al álgebra de Lie de todos los campos vectoriales invariantes por la izquierda.

El álgebra de Lie de un grupo de Lie refleja las propiedades del grupo y lo hace de una manera simple, pues tiene como sustento un espacio vectorial y no una variedad diferencial; esto la hace sumamente útil para estudiar la estructura y las propiedades del grupo. Es importante notar que esta álgebra de Lie se sustenta en el espacio vectorial T_eG, ello implica que las propiedades del grupo de Lie se relacionan con sus propiedades locales en un entorno de su identidad.

Ahora dado un vector $V \in T_eG$, o sea, un elemento de $\mathfrak{g}$, se genera un campo vectorial X_V y su flujo ψ_V^s, así tendremos una curva o trayectoria única $x(s)$ sobre G tangente al campo X_V y que pasa por el punto $e \in G$, donde naturalmente será tangente al vector V; en esta trayectoria $x(0) = e$ y $x(s_1 + s_2) = x(s_1) * x(s_2)$, esto es, conforma un subgrupo de dimensión uno de G, la curva inmersa en G es la variedad diferencial de este subgrupo y la operación, naturalmente, es la misma que en G. Podemos ahora establecer un mapeo, que llamaremos el mapeo exponencial, entre el álgebra de Lie del grupo y el propio grupo $\exp : \mathfrak{g} \to G$ por medio de

$$\exp(sV) = x(s), \tag{5.178}$$

el teorema de existencia y unicidad nos muestra que este mapeo está bien definido y asocia a la linea sV de $\mathfrak{g}$ la trayectoria o subgrupo de G $x(s)$. El

flujo ψ_V^s actuará en G por medio de

$$\psi_V^s(g) = g * \psi_V^s(e) = g * x(s) = g * \exp(sV). \tag{5.179}$$

Podemos también establecer una correspondencia inversa, dado un subgrupo de un parámetro ψ^s de difeomorfismos de G, esto es, $\psi^0 = e$ y $\psi^{s_1+s_2} = \psi^{s_1} * \psi^{s_2}$, entonces existe un elemento $W \in \mathfrak{g}$ tal que $\psi^s = \exp(sW)$.

En la mecánica clásica adquiere importancia el concepto de la acción de un grupo de Lie sobre una variedad diferencial M, que en muchos casos resuta ser el espacio fase. Una acción diferenciable del grupo de Lie G en la variedad M es un mapeo $\Theta : G \times M \to M$ que dados un elemento del grupo $g \in G$ y un punto p de M genera un punto de M que denotaremos por $\Theta_g(p)$[14] con las siguientes propiedades:

El mapeo Θ es diferenciable.

El mapeo $\Theta_e : M \to M$ es la identidad, esto es $\Theta_e(p) = p \ \forall \ p \in M$.

El mapeo Θ cumple con $\Theta_{g_1} \circ \Theta_{g_2} = \Theta_{g_1 * g_2}$.

Podemos ver que los grupos de un parámetro de difeomorfismos de la variedad de configuración o del espacio fase que hemos utilizado, son casos particulares de las acciones del grupo de Lie R, esto es, la recta real como variedad diferencial y la operación suma de reales como la operación de grupo. Desde luego esta definición de acción simplemente amplia la que hicimos en el primer capítulo para definir los sistemas dinámicos en R^n.

Cuando tenemos la acción de un grupo de Lie sobre un espacio vectorial por medio de mapeos lineales, esto es, tomamos a un espacio vectorial normado como variedad diferencial y los mapeos Θ_g lineales, se dice que la acción es una representación del grupo. Podemos ver que, en ese caso, al tomar una base en el espacio vectorial la acción del grupo puede estudiarse utilizando a las matrices que representan a los mapeos, naturalmente estas matrices formarán un grupo ante la operación de multiplicación isomorfo al grupo de los mapeos Θ_g ante la composición, el cual es, a su vez, homomorfo al grupo de Lie original. En muchas ocasiones a estos grupos de matrices se les llama también representaciones del grupo y proporcionan una forma muy útil de estudiar la estructura y propiedades de los grupos de Lie.

Consideremos una acción $\Theta : G \times M \to M$ del grupo G sobre la variedad M y tomemos un elemento de $\mathfrak{g}$, esto es, un vector V de T_eG. El mapeo exponencial nos permite construir un mapeo de R sobre G por medio de $\exp(sV)$, donde $s \in R$; al combinar este mapeo con Θ podemos obtener una acción del grupo R sobre M al hacer que, a $s \in R$ y a $p \in M$, les

[14] Denotaremos por Θ_g al mapeo de M en M que se obtiene para cada $g \in G$ fija.

corresponda $\Theta(\exp(sV), p) \in M$; está acción es un flujo sobre M y por tanto le corresponde un campo vectorial X_V sobre M con

$$X_V(p) = \frac{d}{ds}\Theta(\exp(sV), p)\Big|_{s=0}, \tag{5.180}$$

a este campo vectorial se le conoce como el generador infinitesimal de la acción Θ correspondiente al vector V. Notamos que en el caso en que la variedad M sea el propio grupo G, el generador infinitesimal es el campo vectorial sobre G definido a partir de D_eL_g por medio de (5.171).

Cuando la variedad M es simpléctica, como ocurre con el espacio fase de un sistema mecánico hamiltoniano, diremos que la acción de un grupo de Lie G sobre M, $\Theta : G \times M \to M$, es una acción canónica si preserva la forma simpléctica, esto es, si para toda $g \in G$, $\Theta_g : M \to M$ es un mapeo diferenciable que preserva la forma simpléctica en M. Notamos que esto es equivalente a la invariancia de los paréntesis de Poisson en M, o sea a

$$\Theta_g[\{F_1, F_2\}] = \{\Theta_g[F_1], \Theta_g[F_2]\} \tag{5.181}$$

donde $F_1, F_2 : M \to R$ y los corchetes indican el paréntesis de Poisson entre las funciones reales de M.

Si Θ es una acción canónica entonces dado un vector del álgebra de Lie $V \in \mathfrak{g}$, el campo vectorial de M o generador infinitesimal X_V dado por (5.180) generará un flujo $\psi_V^s : M \to M$ que deja invariante a la forma simpléctica; por ello el generador infinitesimal X_V será un campo vectorial hamiltoniano, denotaremos por $F_V : M \to R$ a la función hamiltoniana correspondiente.

Ahora (5.180) nos hace ver que la asociación entre los elementos del álgebra de Lie y los generadores infinitesimales es lineal, esto es,

$$X_{\lambda(V_1+V_2)} = \lambda(X_{V_1} + X_{V_2}), \tag{5.182}$$

por tanto si el grupo es de dimensión k finita podemos tomar una base $(e_1, e_2, \ldots, e_k)$ de $\mathfrak{g}$, tendremos que todo vector $V \in \mathfrak{g}$ se podrá poner como

$$V = \sum_{i=1}^{k} V^i e_i \tag{5.183}$$

y entonces

$$X_V = \sum_{i=1}^{k} V^i X_{e_i} \tag{5.184}$$

donde X_{e_i} son los generadores infinitesimales de la acción correspondientes a los vectores de la base. Al ser la acción Θ canónica, todos estos generadores infinitesimales serán campos hamiltonianos con funciones hamiltonianas F_V y F_{e_i} respectivamente. Tendremos entonces que

$$dF_V = \gamma[\bullet, X_V] = \sum_{i=1}^{k} V^i \gamma[\bullet, X_{e_i}] \tag{5.185}$$

$$= \sum_{i=1}^{k} V^i dF_{e_i}, \tag{5.186}$$

vemos entonces que únicamente k de estas funciones hamiltonianas serán linealmente independientes y que las n funciones F_{e_i} forman un conjunto linealmente independiente.

Cuando la variedad M es el espacio fase de un sistema mecánico con hamiltoniana $H : M \to R$ y la acción canónica del grupo G deja invariante a esta hamiltoniana, H será invariante ante el campo X_V para todo $V \to \mathfrak{g}$, o sea

$$X_V[H] = 0, \tag{5.187}$$

pero X_V es un campo hamiltoniano con hamiltoniana F_V, por lo que

$$\{H, F_V\} = X_V[H] = 0 \tag{5.188}$$

y F_V será una constante de movimiento, como hay k de estas funciones linealmente independientes, obtenemos el teorema de Noether para grupos de Lie en general:

Decimos que un sistema mecánico con espacio fase M y función hamiltoniana $H : M \to R$ es simétrico ante un grupo de Lie G de dimensión k, si existe una acción canónica del grupo sobre M que deja invariante a la función hamiltoniana.

Si un sistema mecánico hamiltoniano es simétrico ante un grupo de Lie G de dimensión k, entonces tiene k constantes de movimiento linealmente independientes. Dada una base para el álgebra de Lie del grupo, estas constantes son las funciones de M cuyos campos vectoriales hamiltonianos resulten ser los generadores infinitesimales de la acción correspondientes a cada uno de los vectores de la base.

Tomaremos como un ejemplo de la aplicación de la teoría de los grupos de Lie y el teorema de Noether el caso de las rotaciones en el espacio físico de tres dimensiones y los sistema mecánicos invariantes ante estas rotaciones. Comenzaremos por encuadrar a las rotaciones en el contexto general de las

transformaciones lineales de R^n en R^n. Tomaremos aquí a R^n dotado de un producto interno dado por

$$< x, y >= \sum_i x_i y_i, \tag{5.189}$$

donde $x = (x_1, \ldots, x_n)$ y $y = (y_1, \ldots, y_n)$.

El conjunto de todos los mapeos lineales invertibles (con determinante distinto de cero) de R^n en R^n, al tomar como operación de suma la composición de mapeos, forma un espacio vectorial de dimensión n^2, como ocurre con todo espacio vectorial puede ser visto como variedad diferencial y como grupo ante la operación que lo define y por tanto como grupo de Lie, a este espacio vectorial se le llama el grupo general lineal real, en caso de haber tomado los complejos en vez de los reales tendríamos el grupo general lineal complejo, se les denota por $GL(n, R)$y $GL(n, C)$ respectivamente; el caso complejo es de gran interés para la mecánica cuántica, pero aquí nos quedaremos en el caso real; denotaremos por $I : R^n \to R^n$ a la transformación idéntica.

Entre las transformaciones lineales de R^n en si mismo, aquellas que dejan invariante al producto interno se llaman transformacines ortogonales1. Así una transformación lineal $T : R^n \to R^n$ será ortogonal si y sólo si

$$< T[x], T[y] >=< x, y > \tag{5.190}$$

para toda $x, y \in R^n$. Esto implica que

$$< T[x], T[x] >= (\det(T))^2 < x, x >=< x, x >, \tag{5.191}$$

y por tanto el determinante de una transformación ortogonal es más o menos uno.

Dada la transformación lineal T a la transformación $T^\intercal$ tal que

$$< T[x], y >=< x, T^\intercal[y] >, \tag{5.192}$$

se le llama la transpuesta de T. Vemos que T será ortogonal si y sólo si

$$T^{-1} = T^\intercal. \tag{5.193}$$

La composición de dos transformaciones ortogonales es, claramente, una transformación ortogonal, la identidad es una transformación ortogonal y dada una transformación ortogonal su inversa es su transpuesta que es también ortogonal, por tanto el conjunto de las transformaciones ortogonales forma

un grupo que será un subgrupo de $GL(n,R)$. Ahora las transformaciones ortogonales no forman un espacio vectorial, basta para ver esto señalar que si T es ortogonal y λ es un escalar real, λT no será ortogonal ya que

$$< \lambda T[x], \lambda T[y] >= \lambda^2 < x, y >; \tag{5.194}$$

sin embargo si conforman una variedad diferencial inmersa en $SL(n,R)$. Tomemos el mapeo $M : SL(n,R) \to SL(n,R)$ dado por $T \mapsto T \circ T^{\intercal}$, que manda cualquier transformación de $SL(n,R)$ en una transformación simétrica, tendremos que $M[I] = I$ y $M^{-1}[I]$ será el conjunto de las transformaciones ortogonales, como I es un valor regular de M, $M^{-1}[I]$ será una variedad diferencial inmersa en $GL(N,R)$. Así el conjunto de las transformaciones ortogonales resulta un grupo de Lie, a este grupo de Lie se le llama el grupo ortogonal y se le denota por $O(n)$. Como variedad diferencial $O(n)$ tiene dos componentes ajenas correspondientes a las transformaciones con determinante más uno o menos uno, de ellas la primera forma un subgrupo, la transformación idéntica tiene determinante uno, la composición de dos de ellas es una de ellas y la inversa tendrá también determinante uno, naturalmente la otra componente no formará grupo por si sola. La primera componente será entonces un grupo de Lie, se le conoce como el grupo especial ortogonal, se le denota por $SO(n)$ y su dimensión será la misma que la de $O(n)$, $\frac{1}{2}n(n-1)$.

Las isometrías del espacio euclideano de n dimensiones son las transformaciones de éste que dejan invariante la distancia entre puntos y los ángulos entre rectas, entre ellas están las que dejan fijo un punto, como el espacio euclidiano se puede identificar con R^n dotado de la métrica proporcionada por (5.189), las isometrías serán una acción de $O(n)$ sobre R^n, en especial las rotaciones serán una acción de $SO(n)$ sobre R^n, los elementos de $O(n)$ con determinante negativo constituyen las reflexiones en una acción de $O(n)$ sobre R^n.

Así las rotaciones en el espacio físico son una acción de $SO(3)$. Estas pueden parametrizase de varían formas, una de ellas consiste en utilizar tres rotaciones en torno a tres ejes independientes, digamos el eje x, el y y el z de un sistema cartesiano de coordenadas, tomadas en ese orden, pues recordamos que las rotaciones no conmutan, el orden es convencional pero es sumamente importante ceñirse a él. Otra consiste en utilizar el teorema de Euler que nos dice que toda rotación en R^3 se reduce a una rotación por un ángulo en torno a un eje o dirección, podemos utilizar dos parámetros para determinar el eje y un tercero para señalar el ángulo. Los ángulos de Euler nos proporcionan una tercera opción. Desde luego hay muchísimas otras, en

cada caso podemos utilizar la más conveniente. Estas parametrizaciones de la acción de $SO(3)$sobre R^3 no son otra cosa que coordenadas locales para la variedad que constituye el grupo $SO(3)$, pues el número de parámetros coincide con la dimensión del grupo, de hecho podemos utilizar la segunda para identificar la variedad diferencial $SO(3)$, veamos. Tomemos una esfera solida de dos dimensiones de radio π, un punto de ella nos proporciona una dirección o eje y una distancia al origen, que podemos tomar como el ángulo por el que hay que rotar en torno de dicho eje, así hemos establecido una correspondencia entre los puntos de la esfera sólida y las rotaciones en tres dimensiones, ahora esta correspondencia no es biuníboca ya que es lo mismo rotar por un ángulo π en una dirección que hacerlo en la dirección opuesta, esto es, un punto sobre la superficie de la esfera y el punto diametralmente opuesto representan la misma rotación, entonces para tener un homeomorfismo entre la esfera sólida y el grupo de rotaciones en tres dimensiones debemos identificar cada punto sobre la superficie de la esfera con su antípoda, esto es similar a la construcción de un toro a partir de un rectángulo identificando o pegando los lados opuestos del rectángulo. El espacio que se obtiene es el llamado espacio proyectivo de tres dimensiones, un espacio que no puede sumergirse en R^3; así $SO(3)$ es el espacio proyectivo de tres dimensiones cuando se le ve como variedad diferencial.

Para continuar con nuestro estudio de las rotaciones en tres dimensiones usaremos la parametrización en términos de tres rotaciones sucesivas en torno a los ejes x, y y z, por tres ángulos θ_x, θ_y y θ_z respectivamente, estos ángulos es suficiente tomarlos en el intervalo $[0, \pi]$, pero notamos, por ejemplo que los ángulos $(0, 0, 0)$ y (π, π, π) representan a la misma rotación, la identidad, esto es una consecuencia del hecho de que $SO(3)$ no es homeomorfo a R^3, esta es una más de las pequeñas dificultades que siempre se presentan cuando queremos abarcar toda una variedad o la mayor parte de ella con una sola carta, naturalmente en el caso de los grupos de Lie estas dificultades son poco importantes debido al teorema, demostrado por el propio Lie, que nos dice que la estructura diferencial de un grupo de Lie queda determinada por sus propiedades locales en torno al elemento identidad, esto es, podemos estudiar todo el grupo a partir de su estudio en una carta que comprenda al elemento identidad, así de restrictivas son las condiciones de formar grupo y de que éste genere mapeos diferenciables en la variedad y, desde luego, es de aquí de donde se desprende la utilidad de los grupo de Lie en diversas ramas de la matemática y la física.

Las rotaciones de R^3 en torno al eje x por un ángulo θ_x, están dadas por

la aplicación de la matriz

$$R_{\theta_x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\text{sen}\theta_x \\ 0 & \text{sen}\theta_x & \cos\theta_x \end{pmatrix}, \tag{5.195}$$

con respecto al eje y tendremos

$$R_{\theta_y} = \begin{pmatrix} \cos\theta_y & 0 & \text{sen}\theta_y \\ 0 & 1 & 0 \\ -\text{sen}\theta_y & 0 & \cos\theta_y \end{pmatrix} \tag{5.196}$$

y con respecto al eje z

$$R_{\theta_z} = \begin{pmatrix} \cos\theta_z & -\text{sen}\theta_z & 0 \\ \text{sen}\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}; \tag{5.197}$$

la rotación completa parametrizada por lo ángulos $(\theta_x, \theta_y, \theta_z)$ vendrá dada por la matriz que resulta del producto $R_{\theta_z} R_{\theta_y} R_{\theta_x}$, esto es

$$R_{\theta_x,\theta_y,\theta_z} = \tag{5.198}$$

$$\begin{pmatrix} \cos\theta_y\cos\theta_z - \text{sen}\theta_x\text{sen}\theta_y\text{sen}\theta_z & -\cos\theta_x\text{sen}\theta_z & \text{sen}\theta_y\cos\theta_z + \text{sen}\theta_x\cos\theta_y\text{sen}\theta_z \\ \cos\theta_y\text{sen}\theta_z + \cos\theta_x\text{sen}\theta_y\text{sen}\theta_z & \cos\theta_x\cos\theta_z & \text{sen}\theta_y\text{sen}\theta_z - \text{sen}\theta_x\cos\theta_y\text{sen}\theta_z \\ -\cos\theta_x\text{sen}\theta_y & \text{sen}\theta_x & \cos\theta_x\cos\theta_y \end{pmatrix}. \tag{5.199}$$

Ahora cada una de las rotaciones en torno a uno de los ejes puede verse como un flujo en R^3 donde cada R está representado o parametrizado por el ángulo correspondiente. Cada uno de estos flujos está determinado por un campo vectorial que corresponde a la derivada del flujo con respecto al parámetro evaluada en cada punto, en la notación matricial a este campo vectorial le corresponde la matriz que resulta de derivar con respecto al parámetro la matriz del flujo y evaluar en el valor cero del parámetro. En el caso del eje x el flujo está dado por

$$\mathbf{r}(\theta_\mathbf{x}) = R_{\theta_x}[\mathbf{r}(\mathbf{0})], \tag{5.200}$$

donde $\mathbf{r}$ representa al punto o vector de R^n (x, y, z), el campo vectorial será

$$V_x = A_x[\mathbf{r}] \tag{5.201}$$

donde la matriz A_x esta dada por

$$A_x = \frac{d}{d\theta_x} R_{\theta_x}\Big|_{\theta_x=0}, \tag{5.202}$$

esto es

$$A_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{5.203}$$

y

$$V_x = \begin{pmatrix} 0 \\ -z \\ y \end{pmatrix}, \tag{5.204}$$

similarmente tendremos

$$A_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{5.205}$$

y

$$V_y = \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix}, \tag{5.206}$$

$$A_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{5.207}$$

$$V_z = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}. \tag{5.208}$$

Estos campos vectoriales son los generadores infinitesimales de la acción de $SO(3)$ sobre R^3 y vemos que, matricialmente, son matrices antisimétricas de 3×3, es más vemos que cualquier matriz antisimétrica de tres por tres es una combinación lineal de las matrices A_x, A_y y A_z, por lo que forman una base para el espacio vectorial de las matrices antisimétricas. Estos generadores infinitesimales constituyen una acción del álgebra de Lie del grupo $SO(3)$, por lo que resultan isomorfos al álgebra de Lie conformada por el espacio vectorial de las matrices antisimétricas de tres por tres. En la base dada por A_x, A_y y A_z los conmutadores son

$$[A_x, A_y] = A_z, \tag{5.209}$$

$$[A_z, A_x] = A_y, \tag{5.210}$$

$$[A_y, A_z] = A_x, \tag{5.211}$$

en términos de los campos vectoriales V_x, V_y y V_z tendríamos relaciones totalmente similares. Como las A forman una base, los conmutadores de cualesquiera dos matrices antisimétricas se expresarán como combinaciones de estos conmutadores, es decir, estos conmutadores definen completamente el álgebra de Lie de $SO(3)$. En general los paréntesis de Lie entre los vectores de una base de un álgebra de Lie definen completamente el álgebra y, como son una base, estos paréntesis se expresarán como combinaciones lineales de ellos mismos, a los coeficientes de estas combinaciones lineales se les llama las constantes de estructura del álgebra y en el caso en que se trate del álgebra de un grupo de Lie se les llama también las constantes de estructura del grupo. En el caso de las rotaciones en tres dimensiones tendremos, enumerando por los índice α, β,γ, etc. a x, y y z, que

$$[A_\alpha, A_\beta] = \epsilon_{\alpha\beta\gamma} A_{k\gamma}, \tag{5.212}$$

donde $\epsilon_{\alpha\beta\gamma}$ es uno si los tres índices son una permutación par de (x, y, z), menos uno si son una permutación impar y cero en los demás casos, por ello decimos que $\epsilon_{\alpha\beta\gamma}$ son las constantes de estructura de $SO(3)$. En general el álgebra de Lie de $O(n)$ y de $SO(n)$ es el álgebra del espacio de las matrices antisimétricas de $n \times n$ con el conmutador de matrices como paréntesis de Lie, la dimensión será entonces $n(n-1)/2$, naturalmente sólo en tres dimensiones se da la coincidencia de n con la dimensión del grupo $SO(n)$.

En el caso especial de tres dimensiones a cada matriz antisimétrica de tres por tres podemos asociarle biunívocamente un vector de R^n por medio, por ejemplo, de

$$(a, b, c) \leftrightarrow \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}, \tag{5.213}$$

esto nos genera un isomorfismo de espacio vectorial entre R^3 y el espacio de las matrices antisimétricas. Ahora si aplicamos la matriz a un vector arbitrario $\mathbf{r} = (x, y, z)$ obtenemos

$$\begin{pmatrix} bz - cy \\ cx - az \\ ay - bx \end{pmatrix} \tag{5.214}$$

que identificamos con el producto vectorial $(a, b, c) \times (x, y, z)$. Tomemos dos matrices antisimétricas A y B, les corresponderan dos vectores de R^3 digamos

a y **b**, y tendremos que

$$[A,B][\mathbf{r}] = (AB - BA)[\mathbf{r}] \tag{5.215}$$

$$= A[\mathbf{b}\times\mathbf{r}] - B[\mathbf{a}\times\mathbf{r}] = \mathbf{a}\times(\mathbf{b}\times\mathbf{r}) - \mathbf{b}\times(\mathbf{a}\times\mathbf{r}) \tag{5.216}$$

$$= (\mathbf{a}\times\mathbf{b})\times\mathbf{r}, \tag{5.217}$$

por lo que a la matriz $[A,B]$ le corresponde el vector $\mathbf{a}\times\mathbf{b}$; entonces R^3 con un paréntesis de Lie definido por el producto vectorial resulta ser un álgebra de Lie isomorfa al álgebra de Lie de las matrices antisimétricas y por tanto isomorfa al álgebra de Lie de $SO(3)$.

Tomemos ahora un caso físico, pensemos en un sitema mecánico compuesto de n partículas de masa m_i que interaccionan entre si de tal manera que la hamiltoniana resulta invariante ante el grupo de rotaciones del espacio físico. El espacio fase será R^{2n} y podemos utilizar un sistema canónico de coordenadas $(x_1, y_1, z_1, \ldots, x_n, y_n, z_n, P_{x_1}, p_{y_1}, p_{z_1}, \ldots, p_{x_n}, p_{y_n}, p_{z_n})$, donde (x, y, z) es un sitema cartesiano en el espacio físico.

En el espacio físico tendremos la acción de $SO(3)$ dada por (5.198), esta acción puede subirse al espacio de configuración aplicando (5.171) a cada una de las partículas, en las coordenadas que estamos usando esto sería aplicar una matriz de $3n \times 3n$ con la forma de una matriz hecha de bloques de tres por tres

$$\begin{pmatrix} C & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & C \end{pmatrix}, \tag{5.218}$$

donde C es la matriz (5.123) y todos las bloques de tres por tres fuera de la diagonal son cero. Esta es una acción de $SO(3)$ en el espacio de configuración R^{3n} que está compuesta por una repetición de la acción en R^3; a este tipo de representaciones de grupos de Lie compuestas de otras se les llama representaciones reducibles, a las que no se pueden descomponer en otras, esto es llevarlas a una forma diagonal por bloques en la que cada bloque es una representación del grupo se les llama representaciones irreducibles, éstas constituyen los ladrillos en base a los que podemos construir cualquier otra representación.

Los impulsos p_x, p_y y p_z para la partícula iésima son $m_i\dot{x}_i$, $m_i\dot{y}_i$ y $m_i\dot{z}_i$ por lo que se transformarán ante las rotaciones del espacio físico de la misma manera, esto es, por medio de la matriz (5.198). Esto nos conduce a la acción de $SO(3)$ sobre el espacio fase dada, en las coordenadas canónicas y en la parametrización que estamos usando, por una matriz de la forma (5.218), pero con $2n$ bloques diagonales, esto es una matriz de $6n \times 6n$ diagonal por

bloques de tres por tres dados por (5.198). Por construcción esta acción es canónica, esto es, preserva la forma simpléctica, y, por hipótesis, preserva la función hamiltoniana.

La acción de $SO(3)$ sobre el espacio fase induce una acción del álgebra de Lie del grupo sobre este mismo espacio fase y como la dimensión del grupo es tres tendremos que esta acción del álgebra de lie puede descomponerse en tres componentes, cada una de ellas se ve como un flujo o acción de R sobre el espacio fase, los tres generadores infinitesimales serán los tres campos vectoriales correspondientes a estos flujos. Por construcción y lo ya hecho para el caso de la parametrización que estamos usando, tendremos que los generadores infinitesimales serán los campos

$$V_{\theta_x} = \begin{pmatrix} 0 \\ -z_i \\ y_i \\ 0 \\ -P_{z_i} \\ P_{y_i} \end{pmatrix}, \tag{5.219}$$

$$V_{\theta_y} = \begin{pmatrix} z_i \\ 0 \\ -x_i \\ P_{z_i} \\ 0 \\ -P_{x_i} \end{pmatrix}, \tag{5.220}$$

$$V_{\theta_z} = \begin{pmatrix} -y_i \\ x_i \\ 0 \\ -P_{y_i} \\ P_{x_i} \\ 0 \end{pmatrix}, \tag{5.221}$$

todo ello en la base

$$\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i}, \frac{\partial}{\partial p_{x_i}}, \frac{\partial}{\partial p_{y_i}}, \frac{\partial}{\partial p_{z_i}} \right) \tag{5.222}$$

de los espacios tangentes al espacio fase. Estos tres campos vectoriales deben ser hamiltonianos ya que la acción del grupo es canónica, por tanto existen

tres funciones I_X, I_y y I_z tales que

$$V_{\theta_x} = \begin{pmatrix} 0 \\ -z_i \\ y_i \\ 0 \\ -P_{z_i} \\ P_{y_i} \end{pmatrix} = \begin{pmatrix} \frac{\partial I_x}{\partial p_{x_i}} \\ \frac{\partial I_x}{\partial p_{y_i}} \\ \frac{\partial I_x}{\partial p_{z_i}} \\ -\frac{\partial I_x}{\partial x_i} \\ -\frac{\partial I_x}{\partial y_i} \\ -\frac{\partial I_x}{\partial z_i} \end{pmatrix}, \tag{5.223}$$

$$V_{\theta_y} = \begin{pmatrix} z_i \\ 0 \\ -x_i \\ P_{z_i} \\ 0 \\ -P_{x_i} \end{pmatrix} = \begin{pmatrix} \frac{\partial I_y}{\partial p_{x_i}} \\ \frac{\partial I_y}{\partial p_{y_i}} \\ \frac{\partial I_y}{\partial p_{z_i}} \\ -\frac{\partial I_y}{\partial x_i} \\ -\frac{\partial I_y}{\partial y_i} \\ -\frac{\partial I_y}{\partial z_i} \end{pmatrix}, \tag{5.224}$$

$$V_{\theta_z} = \begin{pmatrix} -y_i \\ x_i \\ 0 \\ -P_{y_i} \\ P_{x_i} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\partial I_z}{\partial p_{x_i}} \\ \frac{\partial I_z}{\partial p_{y_i}} \\ \frac{\partial I_z}{\partial p_{z_i}} \\ -\frac{\partial I_z}{\partial x_i} \\ -\frac{\partial I_z}{\partial y_i} \\ -\frac{\partial I_z}{\partial z_i} \end{pmatrix}, \tag{5.225}$$

por tanto encontramos que

$$I_x = \sum_{i=1}^{n} \left(y_i p_{z_i} - z_i p_{y_i} \right), \tag{5.226}$$

$$I_y = \sum_{i=1}^{n} \left(z_i p_{x_i} - x_i p_{z_i} \right), \tag{5.227}$$

$$I_z = \sum_{i=1}^{n} \left(x_i p_{y_i} - y_i p_{x_i} \right), \tag{5.228}$$

el teorema de Noether nos señala que estas funciones serán constantes de movimiento, en este caso, integrales de movimiento. Podemos ver que estas integrales de movimiento son las componentes del impulso angular total del sistema.

A lo largo de este ejemplo hemos visto como los conceptos y resultados generales en torno a los grupos de Lie y al teorema de Noether se expresan en un caso concreto. Es de destacar la diferencia entre el grupo de Lie, que es una estructura abstracta, y las acciones del grupo sobre las diversas variedades que intervienen en una aplicación concreta, como es el caso del espacio fase en la mecánica clásica o los espacios vectoriales normados en la teoría de representaciones. Esto mismo ocurre con el álgebra de Lie correspondiente, en la teoría general es una estructura abstracta y se expresa en los casos concretos de muy diversas formas, en especial como acciones del álgebra sobre variedades diferenciales. Es importante también tener claridad sobre las diferencias entre los generadores infinitesimales del grupo o del álgebra, los generadores de las acciones que tengamos y las funciones del espacio fase que resultan ser constantes de movimiento, estas últimas son funciones hamiltonianas correspondientes a los generadores infinitesimales de las acciones sobre el espacio fase, que son campos vectoriales hamiltonianos; los generadores infinitesimales del grupo son, por ejemplo, los campos dados por las matrices antisimétricas que aparecen en (5.210), que nada tienen del vector de impulso angular físico; son las funciones I_x, I_y y I_z que generan hamiltonianamente a los generadores infinitesimales de las acciones sobre el espacio fase, las que resultan integrales de movimiento y son las componentes del impulso angular total.

5.6. Transformaciones canónicas.

En el espacio fase de un sistema mecánico hemos definido una serie de coordenadas locales especiales, a las que llamamos coordenadas canónicas; por definición en este tipo de coordenadas las ecuaciones de movimiento se expresan en la forma hamiltoniana (5.13) o, lo que es equivalente, la forma simpléctica se expresa en la forma (5.139). Hasta ahora siempre hemos obtenido coordenadas de este tipo adicionando a un sistema de coordenadas locales para la variedad de configuración, $(q^1, \ldots, q^n)$, los impulsos correspondientes $(p_1, \ldots, p_n)$, para obtener $(q^1, \ldots, q^n, p_1, \ldots, p_n)$. Desde la perspectiva de estas coordenadas un punto del espacio fase se describe como un punto de la variedad de configuración y una forma lineal que actúa en el espacio tangente en ese punto. Hay muchos de estos sistemas de coordenadas canónicas, por lo pronto uno por cada sistema local de coordenadas para la variedad de configuración, pero es fácil ver que hay sistemas canónicos de coordenadas, esto es, sistemas en los que las ecuaciones de movimiento toman la forma hamiltoniana, que no se derivan de un sistema de coordenadas

de la variedad de configuración.

Por otra parte, el espacio fase es una variedad diferencial y en ella podemos definir un sin número de sistemas locales de coordenadas. La mayoría de ellos no serán canónicos pues en ellos las ecuaciones de movimiento no tendrán la forma hamiltoniana. Hay que remarcar que este asunto no es un asunto intrínseco a los sistemas mecánicos; la cualidad de ser canónico se refiere al sistema de coordenadas que usemos en el espacio fase y escoger un sistema de coordenadas es una cuestión totalmente arbitraria, nosotros lo escogemos en función de nuestra conveniencia, mientras que las propiedades y características físicas de los sistemas mecánicos no dependen de nosotros ni del sistema de coordenadas que escojamos. En este sentido, lo que debemos tomar como una propiedad de los sistemas mecánicos mismos es la existencia para todos ellos de al menos un sistema de coordenadas canónicas, y esto es una consecuencia directa de la estructura simpléctica inherente a todo sistema mecánico. Así debemos contemplar a la ecuación (5.142) como la que define física y objetivamente las ecuaciones de movimiento y lo hace independientemente del sistema de coordenadas, e incluso de que hayamos definido uno o no lo hayamos hecho. Son las ecuaciones (5.13), que llamamos de Hamilton, las que adquieren una forma especial, muy conveniente, cuando expresamos las ecuaciones físicas (5.142) en un sistema canónico de coordenadas, pero debemos recordar que en otros sistemas la representación en coordenadas de las ecuaciones (5.142) no tendrá esa forma hamiltoniana. Esto es similar a lo que ocurre cuando tomamos la formulación vectorial de la mecánica y expresamos las ecuaciones en términos de las componentes de los vectores en algún sistema de coordenadas, la forma de las ecuaciones resultantes dependerá del sistema de coordenadas que hayamos escogido.

Al aplicar una transformación arbitraria de coordenadas a un sistema de coordenadas canónicas obtendremos, en general, un sistema no canónico, sin embargo existen transformaciones de coordenadas que al aplicarlas a un sistema canónico nos llevan a otro sistema canónico de coordenadas y son, por este hecho, de gran importancia. Por ello diremos que una transformación de coordenadas en el espacio fase es canónica si preserva la forma de las ecuaciones de Hamilton; esto es, una transformación de coordenadas que envíe las coordenadas $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ en las coordenadas $(x^1, \ldots, x^n, y_1, \ldots, y_n)$ será una transformación canónica si, al ser canónicas las coordenadas (q, p), lo son también las (x, y) o viceversa.

Para iniciar nuestro estudio de las transformaciones canónicas estudiaremos aquellas que se obtienen a partir de una transformación de coordenadas en la variedad de configuración, esto es, tomamos una transformación de

coordenadas

$$Q^i = Q^i(q^1, \ldots, q^n, t) \tag{5.229}$$

y generamos los impulsos $(P_1, \ldots, P_n)$ utilizando la definición

$$P_i = \frac{\partial L}{\partial \dot{Q}^i}, \tag{5.230}$$

naturalmente el sistema de coordenadas resultante $(Q^1, \ldots, Q^n, P_1, \ldots, P_n)$ será canónico, por lo que la transformación

$$(q^1, \ldots, q^n, p_1, \ldots, p_n) \to (Q^1, \ldots, Q^n, P_1, \ldots, P_n) \tag{5.231}$$

será canónica. A una transformación de este tipo se le llama una transformación de punto; hemos dejado que la transformación de coordenadas en la variedad de configuración pueda depender del tiempo ya que la formulación lagrangiana es invariante ante estas transformaciones. Veamos estas transformaciones desde el punto de vista hamiltoniano.

Tomemos un sistema mecánico con lagrangiana

$$L = L(q, \dot{q}, t) \tag{5.232}$$

y una transformación de coordenadas en la variedad de configuración

$$Q^i = Q^i(q, t) \tag{5.233}$$

o

$$q^i = q^i(Q, t). \tag{5.234}$$

Esto induce un cambio en las velocidades generalizadas

$$\dot{Q}^i = \sum_j \frac{\partial Q^i}{\partial q^j} \dot{q}^j + \frac{\partial Q^i}{\partial t} \tag{5.235}$$

o

$$\dot{q}^i = \sum \frac{\partial q^i}{\partial Q^j} \dot{Q}^j + \frac{\partial q^i}{\partial t}. \tag{5.236}$$

La lagrangiana en el sistema Q será

$$L(Q, \dot{Q}, t) = L(q(Q, t), \dot{q}(Q, \dot{Q}, t), t), \tag{5.237}$$

por lo que el impulso P_i en el sistema mayúsculo será

$$P_i = \frac{\partial L}{\partial \dot{Q}^i} = \sum_j \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \dot{q}^j}{\partial \dot{Q}^i} = \sum_j p_j \frac{\partial q^j}{\partial Q^i} \tag{5.238}$$

o, en sentido inverso

$$p_i = \sum_j P_j \frac{\partial Q^j}{\partial q^i}, \tag{5.239}$$

esto nos proporciona una transformación de coordenadas completa en el espacio fase.

Pasemos a la formulación hamiltoniana. En el sistema minúsculo tendremos

$$H(q,p,t) = \sum_i p_i \dot{q}^i - L(q,\dot{q},t), \tag{5.240}$$

mientras que en el sistema mayúsculo tendremos

$$H'(Q,P,t) = \sum_i P_i \dot{Q}^i - L(Q,\dot{Q},t), \tag{5.241}$$

naturalmente las ecuaciones de movimiento en el sistema mayúsculo serán

$$\dot{P}_i = \frac{\partial H'}{\partial Q^i}, \tag{5.242}$$

$$\dot{Q}^i = -\frac{\partial H'}{\partial P_i} \tag{5.243}$$

y hemos puesto H' debido a que no necesariamente las dos hamiltonianas serán la misma, veamos esto

$$H'(Q(q,t),P(q,p,t),t) = \sum_i \left(\sum_j p_j \frac{\partial q^j}{\partial Q^i} \right) \left(\sum_k \frac{\partial Q^i}{\partial q^k} \dot{q}^k + \frac{\partial Q^i}{\partial t} \right) - L(q,\dot{q},t) \tag{5.244}$$

$$= \sum_{jk} p_j \left(\sum_i \frac{\partial q^j}{\partial Q^i} \frac{\partial Q^i}{\partial q^k} \right) \dot{q}^k + \sum_j p_j \left(\sum_i \frac{\partial q^j}{\partial Q^i} \frac{\partial Q^i}{\partial t} \right) - L(q,\dot{q},t) \tag{5.245}$$

$$= \sum_j p_j \dot{q}^j + \sum_j p_j \frac{\partial q^j}{\partial t} - L(q,\dot{q},t) \tag{5.246}$$

$$= H(q,p,t) + \sum_j p_j \frac{\partial q^j}{\partial t}, \tag{5.247}$$

esto es,

$$H'(Q,P,t) = H(q(Q,t),p(Q,P,t),t) + \sum_j p_j(Q,P,t) \frac{\partial q^j(Q,t)}{\partial t}; \tag{5.248}$$

vemos que, mientras la lagrangiana se transforma simplemente con el cambio de coordenadas, la hamiltoniana adquiere un término adicional. Así, ante una transformación de punto, en la formulación hamiltoniana se debe cambiar la función hamiltoniana para que las ecuaciones de movimiento queden expresadas en forma hamiltoniana. Notamos que si la transformación de punto no depende explicitamente del tiempo, entonces la hamiltoniana H' es simplemente H, esto es, ante transformaciones independientes del tiempo, la hamiltoniana permanece invariante. En pocas palabras, toda transformación de coordenadas en la variedad de configuración induce una transformación canónica en el espacio fase, las ecuaciones de movimiento conservaran la forma hamiltoniana, pero con una nueva función hamiltoniana dada por (5.248). Regresemos entonces al estudio general de las transformaciones canónicas de coordenadas comenzando con aquellas que no dependen del tiempo.

5.6.1. Transformaciones canónicas independientes del tiempo.

Llamemos $(x^1, \ldots, x^n, y_1, \ldots, y_n)$ a un sistema local de coordenadas, no necesariamente canónico, para el espacio fase[15], tomemos una transformación de coordenadas a un sistema canónico (q, p)

$$\begin{aligned} x^i &= x^i(q^1, \ldots, q^n, p_1, \ldots, p_n), \\ y_i &= y_i(q^1, \ldots, q^n, p_1, \ldots, p_n), \end{aligned}$$

o

$$\begin{aligned} q^i &= q^i(x^1, \ldots, x^n, y_1, \ldots, y_n), \\ p_i &= p_i(x^1, \ldots, x^n, y_1, \ldots, y_n). \end{aligned} \tag{5.249}$$

[15]Hemos conservado la notación de poner el índice arriba para las primeras n coordenadas y el índice abajo para las segundas. Es importante hacer ver que cuando tenemos un sistema cualquiera de coordenadas en el espacio fase, pierde sentido el tener índices arriba o abajo ya que no son coordenadas locales de la variedad de configuración unas y componentes de una forma lineal las otras, por lo que no tendrán un cambio covariante o contravariante ante cambios de coordenadas en la variedad de configuración. En rigor deberíamos denotarlas todas con índice arriba, puesto que su covariancia o contravariancia queda definida únicamente ante transformaciones de coordenadas del propio espacio fase. La notación que hemos escogido nos es útil por recordarnos la que se utiliza en el caso de los sistemas canónicos de coordenadas derivados de un sitema local de coordenadas de la variedad de configuración, por lo que la conservaremos, sin dejar de mencionar que ha dejado de tener sentido el hablar de covariancia o contravariancia ante transformaciones de coordenadas en la variedad de configuración.

Ante este cambio de coordenadas la forma simpléctica se escribe como

$$\gamma = \sum_i dp_i \wedge dq^i = \sum_i \left[\sum_j \left(\frac{\partial p_i}{\partial y_j} dy_j + \frac{\partial p_i}{\partial x^j} dx^j \right) \wedge \sum_k \left(\frac{\partial q^i}{\partial y_k} dy_k + \frac{\partial q^i}{\partial x^k} dx^k \right) \right] \tag{5.250}$$

$$= \sum_{j<k} \left[\sum_i \left(\frac{\partial p_i}{\partial x^j} \frac{\partial q^i}{\partial x^k} - \frac{\partial p_i}{\partial x^k} \frac{\partial q^i}{\partial x^j} \right) dx^j \wedge dx^k \right] \tag{5.251}$$

$$+ \sum_{j<k} \left[\sum_i \left(\frac{\partial p_i}{\partial y_j} \frac{\partial q^i}{\partial y_k} - \frac{\partial p_i}{\partial y_k} \frac{\partial q^i}{\partial y_j} \right) \right] dy_j \wedge dy_k \tag{5.252}$$

$$+ \sum_{jk} \left[\sum_i \left(\frac{\partial p_i}{\partial y_j} \frac{\partial q^i}{\partial x^k} - \frac{\partial p_i}{\partial x_k} \frac{\partial q^i}{\partial y^j} \right) \right] dy_j \wedge dx^k, \tag{5.253}$$

que no se reduce, en general, a $\sum dy_i \wedge dx^i$. Esto es la forma simpléctica no queda invariante ante el cambio de coordenadas y por tanto, como las ecuaciones de Hamilton se derivan de (5.142) al expresar la forma simpléctica y la diferencial de la función hamiltoniana en coordenadas, la forma de las ecuaciones de movimiento no quedará hamiltoniana. Hemos encontrado que, ante un cambio general de coordenadas en el espacio fase, las ecuaciones de Hamilton no son covariantes[16].

Ahora bien, como hemos visto, un cambio de coordenadas de la variedad de configuración induce una transformación canónica de coordenadas en el espacio fase que, naturalmente deja invariante a la forma simpléctica. Para comprobar esto basta ver que un cambio de coordenadas en la variedad de configuración

$$q^i = q^i(x^1, \ldots, x^n), \tag{5.254}$$

[16]La formulación lagrangiana es covariante ante cambios de coordenadas en la variedad de configuración y no ante cambios de coordenadas generales en el haz tangente, esto es, la covariancia se da cuando pasamos de coordenadas $(q, \dot{q})$ a $(Q, \dot{Q})$, que no es un cambio general en el haz tangente. Naturalmente en la formulación lagrangiana no expresamos en general las ecuaciones de movimiento en el haz tangente, esto es como sistemas de ecuaciones de primer orden, por lo que no se contemplan como importantes los cambios generales de coordenadas y siempre nos limitamos a los cambios de coordenadas generalizadas y los cambios que estos inducen en las velocidades generalizadas. Por el contrario en la formulación hamiltoniana son importantes los cambios de coordenadas en el espacio fase de $2n$ dimensiones, por lo que el asunto de la covariancia se presenta de manera mucho más amplia.

se completa a un cambio en el espacio fase por

$$p_l = \sum_j y_j \frac{\partial x^j}{\partial q^l}, \tag{5.255}$$

y al substituir en (5.253) obtenemos la invariancia de la forma simpléctica.

En el espacio fase pueden existir transformaciones de coordenadas más generales que dejan invariante la forma simpléctica, o, lo que es equivalente, la forma de las ecuaciones de Hamilton. De las ecuaciones (5.253) deducimos que son todas aquellas para las que se cumplan las condiciones

$$\begin{aligned}
\sum_i \left(\frac{\partial p_i}{\partial x^j} \frac{\partial q^i}{\partial x^k} - \frac{\partial p_i}{\partial x^k} \frac{\partial q^i}{\partial x^j} \right) &= 0, \\
\sum_i \left(\frac{\partial p_i}{\partial y_j} \frac{\partial q^i}{\partial y_k} - \frac{\partial p_i}{\partial y_k} \frac{\partial q^i}{\partial y_j} \right) &= 0, \\
\sum_i \left(\frac{\partial p_i}{\partial y_j} \frac{\partial q^i}{\partial x^k} - \frac{\partial p_i}{\partial x^k} \frac{\partial q^i}{\partial y_j} \right) &= \delta^j_k,
\end{aligned} \tag{5.256}$$

pues de ser así tendremos que

$$\sum_i dp_i \wedge dq^i = \sum_i dy_i \wedge dx^i. \tag{5.257}$$

Cuando tenemos que las coordenadas canónicas (q, p) son función de dos o más argumentos $u, v, w, \ldots$ se definen los paréntesis de Lagrange por medio de

$$[u, v] \equiv \sum_i \left(\frac{\partial q^i}{\partial u} \frac{\partial p_i}{\partial v} - \frac{\partial q^i}{\partial v} \frac{\partial p_i}{\partial u} \right), \tag{5.258}$$

que no son otra cosa que la suma de los determinantes jacobianos

$$\left| \frac{\partial (q^i, p_i)}{\partial (u, v)} \right|. \tag{5.259}$$

Cuando tenemos las transformaciones de coordenadas (5.249), podemos ver a las coordenadas canónicas como funciones de las coordenadas (x, y), por ello podemos expresar las condiciones de canonicidad (5.256) por medio de

$$\left[x^k, x^j \right] = 0, \tag{5.260}$$

$$[y_k, y_j] = 0, \tag{5.261}$$

$$\left[x^k, y_j\right] = \delta_j^k. \tag{5.262}$$

Cuando las coordenadas canónicas dependen de $2n$ funciones, digamos u_α, $\alpha = 1, \ldots, 2n$, de manera linealmente independiente, se establece una relación entre los paréntesis de Lagrange y los de Poisson

$$\sum_\alpha [u_\alpha, u_\beta]\{u_\alpha, u_\gamma\} = \delta_{\beta\gamma}, \tag{5.263}$$

para demostrar esto basta substituir los paréntesis por sus definiciones y hacer el cálculo directamente. Usando esta relación podemos, de manera totalmente equivalente a (5.261), expresar las condiciones de canonicidad por

$$\left\{x^k, x^l\right\}_{(q,p)} = 0, \tag{5.264}$$

$$\{y_k, y_l\}_{(q,p)} = 0, \tag{5.265}$$

$$\left\{x^k, y_l\right\}_{(q,p)} = \delta_l^k, \tag{5.266}$$

donde hemos puesto la indicación que los paréntesis deben ser calculados en las coordenadas (q, p).

Tenemos así varias formas totalmente equivalentes de definir una transformación canónica de coordenadas:

1. La que deja invariante la forma de las ecuaciones de Hamilton.
2. La que deja invariante la expresión de la forma simpléctica, ver ecuación (5.257).
3. La que, entendida como un difeomorfismo del espacio fase en si mismo, deja invariante a la forma simpléctica.
4. La que cumple con las relaciones (5.261).
5. La que cumple con las relaciones (5.265).

Todas ellas pueden verse como definición o tomarse como criterio para analizar si una transformación dada es canónica o no lo es. Debemos recordar que la invariancia señalada en uno y dos se refiere a la de la forma que adquieren las ecuaciones o expresiones en coordenadas canónicas y que los paréntesis de Lagrange y de Poisson se calculan en las coordenadas originales suponiéndolas canónicas.

Es costumbre denotar por (q,p) a un sistema canónico de coordenadas, como hemos venido haciendo, y por

$$Q^i = Q^i(q,p), \tag{5.267}$$

$$P_i = P_i(q,p), \tag{5.268}$$

a una transformación canónica de coordenadas, esto independientemente de que unas representen unas coordenadas generalizadas para la variedad de configuración y las otras impulsos generalizados o no. En adelante usaremos esta convención a no ser que se indique otra cosa. Veamos algunos ejemplos de transformaciones canónicas de coordenadas.

La transformación

$$Q^i = p_i, \tag{5.269}$$

$$P_i = -q^i, \tag{5.270}$$

resulta ser canónica, como podemos hace ver simplemente utilizando los criterios (5.265). Notamos que hemos intercambiado coordenadas por impulsos, esto nos hace ver de una manera muy vivida que las coordenadas no contienen la estructura del espacio fase, ni siquiera las coordenadas canónicas, simplemente la estructura se refleja de manera distinta en los distintos sistemas de coordenadas.

Pensemos ahora en un oscilador armónico. En las coordenadas habituales su hamiltoniana es

$$H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2, \tag{5.271}$$

donde (x,p) forman un sitema canónico de coordenadas. Las ecuaciones de movimiento son

$$\dot{q} = \frac{p}{m}, \tag{5.272}$$

$$\dot{p} = -kx. \tag{5.273}$$

Sabemos que en el espacio fase las trayectorias son curvas cerradas concéntricas, esto nos llama a buscar un sistema de coordenadas en el espacio fase del tipo de las coordenadas polares y que resulten canónicas, proponemos así una transformación de coordenadas a coordenadas (θ, I) por medio de

$$x = a\sqrt{I}\text{sen}\theta, \tag{5.274}$$

$$p = b\sqrt{I}\cos\theta. \tag{5.275}$$

Para que esta transformación sea canónica los criterios (5.265) nos indican que

$$\{x,p\} = \frac{ab}{2}\left(\text{sen}^2\theta + \cos^2\theta\right) = 1, \tag{5.276}$$

por tanto $b = 2/a$ y la transformación canónica será

$$x = a\sqrt{I}\text{sen}\theta, \tag{5.277}$$

$$p = \frac{2}{a}\sqrt{I}\cos\theta. \tag{5.278}$$

La hamiltoniana en estas nuevas coordenadas será (5.271) expresada en (θ, I), ya que la transformación no depende del tiempo, entonces

$$H = \frac{2}{ma^2} I\cos^2\theta + \frac{ka^2}{2} I\text{sen}^2\theta, \tag{5.279}$$

como a es arbitraria podemos tomar el valor

$$a^2 = \frac{2}{\sqrt{km}} \tag{5.280}$$

y obtenemos que la hamiltoniana se expresa como

$$H = \sqrt{\frac{k}{m}} I; \tag{5.281}$$

observamos que I es una integral de movimiento por ser θ ignorable y las ecuaciones de movimiento, que serán hamiltonianas, serán

$$\dot{\theta} = \sqrt{\frac{k}{m}}, \tag{5.282}$$

$$\dot{I} = 0, \tag{5.283}$$

con soluciones

$$\theta = \sqrt{\frac{k}{m}} t + \beta \tag{5.284}$$

y

$$I = C. \tag{5.285}$$

Vemos la enorme simplificación de las ecuaciones de movimiento a que nos llevó la transformación de coordenadas. Si deseamos la solución en las coordenadas (x, p) basta aplicar la transformación canónica de coordenadas y obtendremos

$$x = A\text{sen}\left(\sqrt{\frac{k}{m}} t + \beta\right) \tag{5.286}$$

Y

$$p = A\sqrt{km}\cos\left(\sqrt{\frac{k}{m}} t + \beta\right), \tag{5.287}$$

donde hemos definido

$$A^2 \equiv \frac{2C}{\sqrt{km}}. \tag{5.288}$$

La busqueda de transformaciones canónicas puede ser, como hemos visto, de gran utilidad en la simplificación de las ecuaciones de movimiento y, por tanto, en la solución de los sistemas mecánicos. Por otro lado, dado que entre todas las posibles transformaciones de coordenadas las canónicas son muy pocas, puede resultar difícil encontrar este tipo de transformaciones. Sin embargo existe un procedimiento por medio del cual se puede encontrar una transformación canónica de coordenadas asociada a cada función de R^{2n} en los reales, donde R^{2n} es un espacio coordenado para el espacio fase. Veamos esto con detenimiento.

Una transformación es canónica si deja invariante a la forma simpléctica. Esto es

$$Q^i = Q^i(q,p), \tag{5.289}$$

$$P_i = P_i(q,p), \tag{5.290}$$

será canónica si

$$\gamma = \sum_i dp_i \wedge dq^i = \sum_i dP_i \wedge dQ^i. \tag{5.291}$$

Tomemos las uno formas diferenciales

$$\omega = \sum_i p_i dq^i + df \tag{5.292}$$

y

$$\Omega = \sum_i P_i dQ^i + df', \tag{5.293}$$

donde f y f' son funciones del espacio fase en los reales. Si la transformación fuese canónica tanto la diferencial de ω como la de Ω serían la forma simpléctica, por lo que la diferencia entre ellas debería ser una diferencial exacta; esto es si la transformación es canónica

$$\sum_i p_i dq^i - \sum_i P_i dQ^i = dF. \tag{5.294}$$

La función F se llama función generadora de la transformación canónica, ya que, si la tenemos, podemos obtener la transformación por medio del siguiente procedimiento.

Dada una función F del espacio fase en los reales, la podremos describir utilizando como coordenadas a las q y las Q si se cumple que el determinante

$$\left|\frac{\partial^2 F}{\partial q^i \partial Q^j}\right| \neq 0, \tag{5.295}$$

esto se debe a que podemos despejar y poner a F como función de las q y de las p o de las Q y las P. La diferencial de esta función será

$$dF = \sum_i \frac{\partial F}{\partial q^i} dq^i + \sum_i \frac{\partial F}{\partial Q^i} dQ^i, \tag{5.296}$$

por lo que, usando (5.294), tendremos que la transformación (5.290) será canónica si

$$\begin{aligned} p_i &= \frac{\partial F}{\partial q^i}, \\ P_i &= -\frac{\partial F}{\partial Q^i}; \end{aligned} \tag{5.297}$$

estas fórmulas pueden ahora ser reacomodadas para obtener la transformación canónica

$$Q^i = Q^i(q,p), \tag{5.298}$$

$$P_i = P_i(q,p). \tag{5.299}$$

Cuando esto ocurra diremos que la transformación canónica ha sido generada por la función F. Es posible también el proceso inverso, esto es, dada una transformación canónica encontrar una función que la genere, basta para ello integrar las ecuaciones (5.297), naturalmente esto será posible únicamente si la transformación es tal que se cumpla (5.295).

Ahora bien es deseable tener una función generadora asociada a toda transformación generadora, el tener que cumplir con (5.295) aparentemente nos lo impide, pero podemos expresar la función F, que desde luego existe para toda transformación canónica, en términos de otras combinaciones de las coordenadas originales y de las transformadas y cumplir una condición análoga a (5.295), veamos esto.

Pensemos en una función F que describimos en términos de las q y de las P y que cumple con

$$\left|\frac{\partial^2 F}{\partial q^i \partial P_j}\right| \neq 0. \tag{5.300}$$

Tomemos la ecuación (5.294) y restemos de ambos términos la cantidad

$$\sum_i Q^i dP_i, \tag{5.301}$$

obtenemos

$$dF - \sum_i Q^i dP_i = \sum_i p_i dq^i - \sum_i P_i dQ^i - \sum_i Q^i dP_i = \sum_i p_i dq^i - d\left(\sum_i Q^i P_i\right), \tag{5.302}$$

por lo que, denotando por F' a $F + \sum_i Q^i P_i$, tendremos

$$dF' = \sum_i p_i dq^i + \sum_i Q^i dP_i. \tag{5.303}$$

Ahora al expresar F' en términos de las q y las P tendremos como condición de canonicidad

$$dF' = \sum_i \frac{\partial F'}{\partial q^i} dq^i + \sum_i \frac{\partial F'}{\partial P_i} dP_i = \sum_i p_i dq^i + \sum_i Q^i dP_i, \tag{5.304}$$

esto es,

$$\begin{aligned} p_i &= \frac{\partial F'}{\partial q^i}, \\ Q^i &= \frac{\partial F'}{\partial P_i}, \end{aligned} \tag{5.305}$$

al reacomodar estas ecuaciones obtendremos una transformación canónica; diremos entonces que la transformación fue generada por la función F'. En sentido inverso dada una transformación canónica, por medio de (5.305) podemos encontrarle una función generadora, claro, será valida, únicamente si ocurre que se cumple (5.300). Este procedimiento es la aplicación de lo que se conoce como una transformación de Legendre, que nos es útil para cambiar las variables en que expresamos alguna condición como (5.294).

De la misma manera podemos hacer que la función generadora de transformaciones canónicas se exprese como función de (Q, p) o de (P, p) en estos casos se deberá cumplir con condiciones análogas a (5.295) o (5.300). Se suele decir que la función generadora es de tipo uno, dos, tres o cuatro cuando se expresa en las diferentes combinaciones de coordenadas minúsculas y mayúsculas; como cualquier función se podrá expresar en alguna de estas combinaciones y cumpliendo con la condición correspondiente, toda transformación canónica será generada por, al menos, una de estas funciones, esto es, entre los cuatro tipos se cubren todas las posibilidades.

En total tendremos cuatro formas para describir en coordenadas una función generadora de transformaciones canónicas, a cada una de ellas corresponde una fórmula para encontrar la transformación a partir de la función

y una condición de validez del procedimiento, según la siguiente tabla

$$F_1(q,Q) \qquad p_i = \frac{\partial F_1}{\partial q^i} \qquad P_i = -\frac{\partial F_1}{\partial Q^i} \qquad \left|\frac{\partial^2 F_1}{\partial q^i \partial Q^j}\right| \neq 0 \tag{5.306}$$

$$F_2(q,P) \qquad p_i = \frac{\partial F_2}{\partial q^i} \qquad Q^i = \frac{\partial F_2}{\partial P_i} \qquad \left|\frac{\partial^2 F_2}{\partial q^i \partial P_i}\right| \neq 0 \tag{5.307}$$

$$F_3(p,Q) \qquad q_i = -\frac{\partial F_3}{\partial p^i} \qquad P_i = \frac{\partial F_3}{\partial Q^i} \qquad \left|\frac{\partial^2 F_3}{\partial p_i \partial Q^i}\right| \neq 0 \tag{5.308}$$

$$F_4(p,P) \qquad q^i = -\frac{\partial F_4}{\partial p_i} \qquad Q^i = -\frac{\partial F_4}{\partial P_i} \qquad \left|\frac{\partial^2 F_4}{\partial p_i \partial P_i}\right| \neq 0 \tag{5.309}$$

Dos funciones generadoras son singularmente significativas, la que genera la transformación idéntica y aquella que genera la transformación (5.270) que intercambia las coordenadas por los impulsos; estas funciones generadoras son

$$F_2(q,P) = \sum_i q^i P_i \tag{5.310}$$

y

$$F_1(q,Q) = \sum_i q^i Q^i \tag{5.311}$$

respectivamente. La primera la expresamos en términos del tipo dos y la segunda del tipo uno.

En (5.278) tenemos una transformación canónica que resultó útil para simplificar la solución del oscilador armónico, veamos cual es su función generadora. Tomemos una función F_1 de tipo uno, que en este caso quiere decir que se considera función de x y θ, entonces la transformación debe estar dada por

$$p = \frac{\partial F_1}{\partial x} \tag{5.312}$$

y

$$I = -\frac{\partial F_1}{\partial \theta}, \tag{5.313}$$

pero de (5.278) tenemos que

$$I = \frac{x^2}{a^2 \mathrm{sen}^2\theta} \tag{5.314}$$

y

$$p = \frac{2}{a}\sqrt{I}\cos\theta = \frac{2}{a^2}x\frac{\cos\theta}{\mathrm{sen}\theta}, \tag{5.315}$$

por tanto debemos integrar las ecuaciones

$$\frac{\partial F_1}{\partial x} = \frac{2}{a^2} x \frac{\cos\theta}{\operatorname{sen}\theta}, \tag{5.316}$$

$$\frac{\partial F_1}{\partial \theta} = -\frac{x^2}{a^2 \operatorname{sen}^2\theta}. \tag{5.317}$$

De la primera de estas ecuaciones vemos que

$$F_1 = \frac{x^2}{a^2} \frac{\cos\theta}{\operatorname{sen}\theta} + g(\theta), \tag{5.318}$$

donde $g(\theta)$ es una función arbitraria de θ. Derivando con respecto a θ vemos que para que (5.318) sea la solución que buscamos $g(\theta)$ debe ser una constante que podemos considerar vale cero, esto es la función generadora es

$$F_1(x, \theta) = \frac{x^2}{a^2} \frac{\cos\theta}{\operatorname{sen}\theta}. \tag{5.319}$$

Es fácil hacer ver que la condición (5.295) se cumple, por lo que tenemos una buena función generadora.

5.6.2. Transformaciones canónicas dependientes del tiempo.

Como hemos visto cuando una transformación canónica depende del tiempo la función hamiltoniana cambia, esto es, en el nuevo sistema de coordenadas debemos utilizar una nueva función hamiltoniana para que las ecuaciones de movimiento del sistema mecánico queden expresadas en la forma hamiltoniana; debemos recordar que la definición y los criterios para que una transformación sea canónica son los mismos ya que no tienen nada que ver con la función hamiltoniana, simplemente, si la transformación depende del tiempo, se deben cumplir para todos los valores del tiempo.

Cuando la transformación de coordenadas depende del tiempo no se puede usar el procedimiento anterior para encontrar una función generadora, esto se debe a que la función dependería a del tiempo y su diferencial contendría un término adicional de la forma

$$\frac{\partial F}{\partial t} dt, \tag{5.320}$$

no podríamos, entonces, utilizar las ecuaciones (5.294) y (5.296) para encontrar la (5.297).

Algo similar ocurre con la formulación hamiltoniana misma al utilizar la forma simpléctica y la relación (5.142), pues la diferencial de una hamiltoniana dependiente del tiempo tendría un término en la diferencial del tiempo; por esta razón en las secciones anteriores nos limitamos al caso en que la hamiltoniana no depende explícitamente del tiempo.

Para ampliar nuestros estudios al caso en que hay dependencia explicita del tiempo, ya sea en la hamiltoniana, la transformación o ambas, debemos adicionar una dimensión temporal al espacio fase y construir la variedad diferencial $TM^* \times R$ de $2n+1$ dimensiones en donde la nueva dimensión es el tiempo, a esta variedad la llamaremos espacio fase extendido. En este espacio fase extendido las trayectorias del sistema mecánico serán curvas cuya proyección a una subvariedad $t = cst.$ serán las trayectorias ordinarias en el espacio fase, pero naturalmente, se irán extendiendo uniformemente a lo largo de la dimensión temporal. Esto es, en el espacio fase extendido las trayectorias conformarán el flujo de un campo vectorial con componente uno a lo largo de la dimensión temporal y que en las demás dimensiones es hamiltoniano, en coordenadas canónicas (q,p,t)[17] este campo se expresará por

$$V_H = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} \right) + \frac{\partial}{\partial t}. \tag{5.321}$$

En el espacio fase extendido no se puede construir una forma simpléctica pues es de dimensión impar. Esto se debe a que en una variedad de dimensión impar toda dos forma diferencial es degenerada, esto es, existe un campo vectorial tal que al aplicar la dos forma diferencial a ese campo y a cualquier otro el resultado es cero. Esto es consecuencia de que en un espacio vectorial de dimensión impar toda dos forma lineal es degenerada, veamos esto. En una base arbitraria la forma se expresará por una matriz antisimétrica $\omega_{ij} = -\omega_{ji}$ y como el determinante de toda matriz es igual al de su transpuesta tendremos

$$|\omega_{ij}| = |\omega_{ji}| = |-\omega_{ij}| = (-1)^m |\omega_{ij}|, \tag{5.322}$$

donde m es la dimensión del espacio vectorial, por lo que, si es impar, el determinante es cero y la matriz tiene un eigenvalor cero, el eigenvector correspondiente será tal que al aplicar ω a ese eigenvector y a cualquier otro vector resultará cero. En el caso de una variedad y una dos forma diferencial

[17] En el espacio fase extendido diremos que un sistema de coordenadas es canónico si tiene la forma (q,p,t) donde la primera parte (q,p) formaría un sistema canónico de coordenadas en el espacio fase y la t es el tiempo, esto es posible dado que el espacio fase extendido es $TM^* \times R$.

eso ocurrirá en cada espacio tangente, por lo que el campo vectorial constituido por los eigenvectores correspondientes en cada espacio tangente tendrá esa propiedad con respecto a la dos forma diferencial. Esto implica que en el espacio fase extendido no se pueden construir formas simplécticas y en él la ecuación (5.142) no tiene sentido.

Para resolver esta dificultad recordamos que para construir la forma simpléctica partimos de la uno forma diferencial

$$\sum_i p_i dq^i, \tag{5.323}$$

en el espacio fase. Tomemos ahora en el espacio fase extendido la uno forma diferencial

$$\omega = \sum_i p_i dq^i - H dt, \tag{5.324}$$

donde $H(q,p,t)$ es la función hamiltoniana, que puede ahora depender explícitamente del tiempo. A esta uno forma se le llama el invariante diferencial de Poincaré Cartán.

A partir de ella y en forma análoga a como construimos la forma simpléctica, construimos una dos forma diferencial en el espacio fase extendido

$$\Gamma = d\omega = \sum_i dp_i \wedge dq^i - dH \wedge dt \tag{5.325}$$

$$= \sum_i dp_i \wedge dq^i - \sum_i \frac{\partial H}{\partial q^i} dq^i \wedge dt - \sum_i \frac{\partial H}{\partial p_i} dp_i \wedge dt, \tag{5.326}$$

ya que la dimensión del espacio fase extendido es impar, está no será simpléctica por ser degenerada, pero por esto mismo habrá un campo vectorial V tal que, para cualquier otro campo vectorial X,

$$\Gamma[V,X] = 0, \tag{5.327}$$

veamos cual es ese campo V.

Para un campo X arbitrario tendremos que en un sistema canónico de coordenadas (q,p,t)

$$\Gamma[V,X] = \sum_i \begin{vmatrix} dp_i[V] & dq^i[V] \\ dp_i[X] & dq^i[X] \end{vmatrix} - \begin{vmatrix} dH[V] & dt[V] \\ dH[X] & dt[X] \end{vmatrix} = \tag{5.328}$$

$$\sum_i \begin{vmatrix} V_i & V^i \\ X_i & X^i \end{vmatrix} - \begin{vmatrix} \sum_i \left(\frac{\partial H}{\partial q^i} V^i + \frac{\partial H}{\partial p_i} V_i \right) + \frac{\partial H}{dt} V_t & V_t \\ \sum_i \left(\frac{\partial H}{\partial q^i} X^i + \frac{\partial H}{\partial p_i} X_i \right) + \frac{\partial H}{dt} X_t & X_t \end{vmatrix}, \tag{5.329}$$

agrupando términos en las componentes del campo X obtenemos

$$\Gamma[V,X]=\sum_i\left[\left(V_i+\frac{\partial H}{\partial q^i}V_t\right)X^i-\left(V^i-\frac{\partial H}{\partial p_i}V_t\right)X_i-\left(\frac{\partial H}{\partial q^i}V^i+\frac{\partial H}{\partial p_i}V_i\right)X_t\right], \tag{5.330}$$

para que se cumpla que $\Gamma[V,X]=0$ para toda X tendremos que deben ser cero las cantidades entre paréntesis que multiplican a las componentes de X, esto es

$$V_i=-\frac{\partial H}{\partial q^i}V_t, \tag{5.331}$$

$$V^i=\frac{\partial H}{\partial p_i}V_t, \tag{5.332}$$

las que al cumplirse hacen que la cantidad que multiplica a X_t también se haga cero, por lo que V_t puede tomar cualquier valor, en particular uno, en cuyo caso

$$V_i=-\frac{\partial H}{\partial q^i}, \tag{5.333}$$

$$V^i=\frac{\partial H}{\partial p_i}, \tag{5.334}$$

$$V_t=1, \tag{5.335}$$

que son las componentes del campo vectorial hamiltoniano en coordenadas canónicas, vemos por tanto que el campo que hace que $\Gamma[V,\bullet]$ sea una unoforma diferencial nula es el campo vectorial hamiltoniano. Esta forma hamiltoniana de expresar las leyes de la mecánica es aplicable cuando la función hamiltoniana depende explícitamente del tiempo y es totalmente equivalente a la expresión (5.142) cuando no hay esa dependencia.

En el caso del estudio de las transformaciones canónicas dependientes del tiempo y sus funciones generadoras las dificultades son similares por lo que podemos utilizar procedimientos semejantes.

Una transformación de coordenadas en el espacio fase extendido

$$\begin{aligned} Q^i &= Q^i(q,p,t),\\ P_i &= P_i(q,p,t),\\ T &= T(q,p,t), \end{aligned}$$

será canónica si al ser (q,p,t) un sistema de coordenadas canónicas, (Q,P,T) también lo es. Esto implica que el tiempo no se transforma $T=t$ y que limitadas al espacio fase (Q,P) son coordenadas canónicas para todo valor del tiempo. En ese caso si el campo vectorial hamiltoniano en el espacio

extendido tiene la expresión (5.321) en las coordenadas minúsculas, en las mayúsculas tendrá la forma

$$\sum_{i=1}^{n}\left(\frac{\partial H'}{\partial P_i}\frac{\partial}{\partial Q^i}-\frac{\partial H'}{\partial Q^i}\frac{\partial}{\partial P_i}\right)+\frac{\partial}{\partial t}. \tag{5.336}$$

Nótese que hemos puesto una nueva hamiltoniana o hamiltoniana transformada H', pues, como hemos visto, esto es indispensable cuando la transformación depende del tiempo.

La forma del campo vectorial hamiltoniano (5.321) se obtiene de la uno forma diferencial (5.324), por lo que para obtener el campo (5.336) debemos partir de la uno forma diferencial

$$\Omega=\sum_i P_i dQ^i - H'dt, \tag{5.337}$$

y para que esos campos sean exactamente el mismo, pero expresado en distintas coordenadas, las diferenciales de ω y de Ω deben tener, en cada espacio tangente, el mismo eigenvector con eigenvalor cero, por lo que deben ser iguales. Para que esto suceda, la diferencia entre ω y Ω debe ser la diferencial de una función del espacio fase extendido que llamaremos $F: T^*M\times R\to R$, esto es

$$\sum_i p_i dq^i - Hdt - \left(\sum_i P_i dQ^i - H'dt\right) = dF. \tag{5.338}$$

Ahora si la función F la expresamos como función de (q,Q,t) suponiendo, para poder hacerlo que

$$\left|\frac{\partial^2 F}{\partial q^i \partial Q^j}\right| \neq 0, \tag{5.339}$$

tendremos que, a todo tiempo

$$\sum_i p_i dq^i - Hdt - \sum_i P_i dQ^i + H'dt = \sum_i\left(\frac{\partial F}{\partial q^i}dq^i+\frac{\partial F}{\partial Q^i}dQ^i\right)+\frac{\partial F}{\partial t}dt, \tag{5.340}$$

por lo que

$$\begin{aligned} p_i &= \frac{\partial F}{\partial q^i}, \\ P_i &= -\frac{\partial F}{\partial Q^i}, \\ H' &= H+\frac{\partial F}{\partial t}. \end{aligned} \tag{5.341}$$

Vemos que dada una función $F(q,Q,t)$ que cumpla con (5.339) por medio de las ecuaciones (5.341) encontramos una transformación canónica y la expresión para la nueva hamiltoniana. Esto es análogo a lo que hicimos en el caso independiente del tiempo, es más las ecuaciones (5.341) se reducen a las (5.297) en el caso en que la función F no dependa explícitamente del tiempo, pues entonces $H' = H$. De la misma manera en que se hizo en el caso independiente del tiempo, podemos a partir de este resultado generar otros similares cuando expresamos la función generadora F en términos de distintas combinaciones de variables minúsculas y mayúsculas, obteniéndose la tabla

$$F_1(q,Q) \quad p_i = \frac{\partial F_1}{\partial q^i} \quad P_i = -\frac{\partial F_1}{\partial Q^i} \quad H' = H + \frac{\partial F_1}{\partial t} \quad \left|\frac{\partial^2 F_1}{\partial q^i \partial Q^j}\right| \neq 0 \tag{5.342}$$

$$F_2(q,P) \quad p_i = \frac{\partial F_2}{\partial q^i} \quad Q^i = \frac{\partial F_2}{\partial P_i} \quad H' = H + \frac{\partial F_2}{\partial t} \quad \left|\frac{\partial^2 F_2}{\partial q^i \partial P_i}\right| \neq 0 \tag{5.343}$$

$$F_3(p,Q) \quad q_i = -\frac{\partial F_3}{\partial p^i} \quad P_i = \frac{\partial F_3}{\partial Q^i} \quad H' = H + \frac{\partial F_3}{\partial t} \quad \left|\frac{\partial^2 F_3}{\partial p_i \partial Q^i}\right| \neq 0 \tag{5.344}$$

$$F_4(p,P) \quad q^i = -\frac{\partial F_4}{\partial p_i} \quad Q^i = -\frac{\partial F_4}{\partial P_i} \quad H' = H + \frac{\partial F_4}{\partial t} \quad \left|\frac{\partial^2 F_4}{\partial p_i \partial P_i}\right| \neq 0. \tag{5.345}$$

Con está tabla están cubiertas todas las posibilidades y a cualquier transformación canónica le corresponde al menos una función generadora y a cada función generadora una transformación canónica.

En la sección anterior encontramos un teorema que establece que un grupo de difeomorfismos del espacio fase de un parámetro deja invariante la forma simpléctica si y solo si está generado de manera hamiltoniana; este teorema se puede poner en términos de transformaciones canónicas, basta para ello recordar que una transformación de coordenadas puede ser vista localmente como un difeomorfismo y viceversa, un difeomorfismo puede interpretarse localmente como una transformación de coordenadas. Así podemos enunciar el teorema diciendo que un grupo de un parámetro de transformaciones de coordenadas será de transformaciones canónicas si y sólo si es generado hamiltonianamente, esto es si sus generadores infinitesimales tiene la forma

$$\left(\frac{\partial G}{\partial p_1}, \ldots, \frac{\partial G}{\partial p_n}, -\frac{\partial G}{\partial q^1}, \ldots, -\frac{\partial G}{\partial q^n}\right), \tag{5.346}$$

donde G es una función de las coordenadas. Dado que estamos formulando el teorema en coordenadas todo es local y los comentarios sobre la localidad del teorema hechos en la sección anterior no se aplican.

Un caso de particular importancia ocurre con el flujo del sistema mecánico, que por ser hamiltoniano, puede verse como un grupo de transformaciones canónicas parametrizado por el tiempo; así podemos identificar la búsqueda de la solución de un sistema hamiltoniano con la búsqueda de ese grupo de transformaciones canónicas o lo que es equivalente la búsqueda de su función generadora.

Lo mismo puede decirse de la expresión general del teorema de Noether en términos hamiltonianos, ya que siempre podemos expresar las acciones canónicas de un grupo de Lie sobre el espacio fase desde la perspectiva de las transformaciones canónicas, esto es, en vez de contemplar la acción del grupo, pensamos en un grupo de transformaciones canónicas. Todo lo dicho en la sección anterior se puede trasladar a la perspectiva de las transformaciones canónicas, claro está en ese caso veremos todo de manera local, ya que las coordenadas son únicamente locales.

Capítulo 6

Formulación de Hamilton Jacobi.

En los capítulos anteriores hemos visto diversas formas de expresar las leyes de la mecánica, desde la formulación Newtoniana o vectorial, hasta la hamiltoniana. Todas ellas son formas distintas de acercarnos a la comprensión del movimiento de los cuerpos, sus causas y sus leyes. Esta multiplicidad de puntos de vista enriquece y profundiza conceptualmente nuestra teoría y genera métodos útiles para atacar el estudio de sistemas mecánicos concretos. Todos ellos tienen una característica común, nos conducen finalmente al establecimiento de un sistema de ecuaciones diferenciales ordinarias, las ecuaciones de movimiento.

En este capítulo veremos otra formulación de la mecánica cuya principal característica es la de no sustentarse directamente en ecuaciones diferenciales ordinarias; en vez de ello tendremos una ecuación diferencial en derivadas parciales, cuyas soluciones, funciones del espacio de configuración y del tiempo, contienen toda la información necesaria para obtener las trayectorias del sistema mecánico; esto la hace radicalmente distinta a las anteriores, nos proporciona un punto de vista más geométrico y nos permite relacionar los fenómenos mecánicos con los del movimiento ondulatorio.

6.1. La ecuación de Hamilton Jacobi.

Consideremos, de manera general, una ecuación diferencial en derivadas parciales de primer orden y en $n+1$ variables

$$Q\left(q^1,\ldots,q^n,t;\frac{\partial G}{\partial q^1},\ldots,\frac{\partial G}{\partial q^n},\frac{\partial G}{\partial t},G\right)=0, \tag{6.1}$$

donde $Q : R^{n+1} \times R^{n+1} \times R \to R$ es una función continua y suficientemente derivable, $G : R^{n+1} \to R$ es una función de las variables en los reales y hemos utilizado la notación de la mecánica $(q^1, \ldots, q^n, t)$ para denotar a las $n+1$ variables. Consideraremos únicamente el caso en que la función Q no depende explícitamente de G. Despejando la derivada de G con respecto a la variable t podemos expresar esta ecuación por medio de

$$K(q^1, \ldots q^n, \frac{\partial G}{\partial q^1}, \ldots, \frac{\partial G}{\partial q^n}, t) = -\frac{\partial G}{\partial t}. \tag{6.2}$$

Se dice que una solución de una ecuación diferencial parcial es general si queda especificada al pedir que, para un valor de alguna de las variables, la solución es una función dada de las demás variables. En nuestra notación es costumbre utilizar al tiempo como la variable distinguida y pedir que la solución G de la ecuación cumpla con

$$G(q^1, \ldots q^n, t = 0) = g(q^1, \ldots, q^n), \tag{6.3}$$

donde g es una función arbitraria dada. Un ejemplo es la conocida solución general $f(x - vt)$ de la ecuación de ondas, la que se reduce a $f(x)$ para $t = 0$.

Otro tipo de soluciones son las llamadas soluciones completas, estas son soluciones que contienen un número máximo de constantes arbitrarias, en nuestro caso por ser la ecuación de primer orden este número máximo es $n+1$, por ser este el número de variables. Por supuesto en la teoría de las ecuaciones diferenciales parciales se contemplan otros tipos de soluciones, como aquellas sujetas a condiciones a la frontera.

Para los propósitos de la teoría de Hamilton Jacobi trataremos únicamente con soluciones completas y soluciones sin condicionamiento específico.

La ecuación (6.2) no contiene explícitamente a la función G, únicamente depende de sus primeras derivadas parciales, por ello una de las constantes de cualquier solución completa será aditiva, como no estamos interesados en constantes aditivas nos quedaremos con soluciones que contengan n constantes no aditivas, a las que llamaremos soluciones completas, en el entendido que únicamente les falta una constante aditiva para serlo plenamente. Denotaremos una solución completa de (6.2) por

$$G(q^1, \ldots q^n, \alpha_1, \ldots, \alpha_n, t), \tag{6.4}$$

esto es, llamamos α_i a las n constantes independientes no aditivas. Pediremos, por razones que quedarán claras más adelante, que adicionalmente las soluciones completas cumplan con la condición

$$\left| \frac{\partial G}{\partial q^i \partial \alpha_j} \right| \neq 0. \tag{6.5}$$

Definamos ahora otras n cantidades, que siguiendo la costumbre de la mecánica denotaremos por p,

$$p_i = p_i(q, \alpha, t) = \frac{\partial G}{\partial q^i}. \tag{6.6}$$

Al substituir en (6.2) las parciales de G por las p definidas en (6.6) obtenemos la identidad

$$K(q^1, \dots q^n, p_1, \dots, p_n, t) \equiv -\frac{\partial G}{\partial t}; \tag{6.7}$$

esto es una identidad y no una ecuación, debido a que hemos establecido que G es solución de (6.2) y que las p se definen por (6.6).

Definamos ahora

$$\beta^i = \frac{\partial G}{\partial \alpha_i}, \tag{6.8}$$

lo que nos permite escribir que

$$\dot{\beta}^i = \frac{d\beta^i}{dt} = \sum_j \frac{\partial^2 G}{\partial q^j \partial \alpha_i}\dot{q}^j + \frac{\partial^2 G}{\partial t \partial \alpha_i}. \tag{6.9}$$

Derivando la identidad (6.7) con respecto a α_i tendremos que

$$\frac{\partial^2 G}{\partial \alpha_i \partial t} = -\sum_j \frac{\partial K}{\partial p_j}\frac{\partial p_j}{\partial \alpha_i} = -\sum_j \frac{\partial K}{\partial p_j}\frac{\partial^2 G}{\partial \alpha_i \partial q^j} \tag{6.10}$$

y por tanto

$$\dot{\beta}_i = \sum_j \left(\dot{q}^j - \frac{\partial K}{\partial p_j}\right)\frac{\partial^2 G}{\partial q^j \partial \alpha_i} \tag{6.11}$$

Por otro lado de la definición de p_i dada por (6.6) tendremos

$$\dot{p}_i = \sum_j \frac{\partial p_i}{\partial q^j}\dot{q}^j + \frac{\partial p_i}{\partial t} = \sum_j \frac{\partial^2 G}{\partial q^j \partial q^i}\dot{q}^j + \frac{\partial^2 G}{\partial t \partial q^i}. \tag{6.12}$$

Diferenciando ahora la identidad (6.7) con respecto a q^i

$$\frac{\partial^2 G}{\partial q^i \partial t} = -\frac{\partial K}{\partial q^i} - \sum_j \frac{\partial K}{\partial p_j}\frac{\partial p_j}{\partial q^i} = -\frac{\partial K}{\partial q^i} - \sum_j \frac{\partial K}{\partial p_j}\frac{\partial^2 G}{\partial q^i \partial q^j}, \tag{6.13}$$

y de aquí obtenemos que

$$\dot{p}_i = \sum_j \left(\dot{q}^j - \frac{\partial K}{\partial p_i}\right)\frac{\partial^2 G}{\partial q^i \partial q^j} - \frac{\partial K}{\partial q^i}. \tag{6.14}$$

Impongamos ahora que las cantidades β sean n constantes arbitrarias con respecto a la variable t, esto es

$$\dot{\beta}_i = 0 = \sum_j \left(\dot{q}^j - \frac{\partial K}{\partial p_j} \right) \frac{\partial^2 G}{\partial q^j \partial \alpha_i}, \tag{6.15}$$

como las segundas parciales de la solución G son independientes tendremos que se debe cumplir con

$$\dot{q}^i = \frac{\partial K}{\partial p_i}, \tag{6.16}$$

y, al cumplirse esto, de (6.14) encontramos que

$$\dot{p}_i = -\frac{\partial K}{\partial q^i}; \tag{6.17}$$

esto es: al considerar, como puede hacerse a partir de la condición de constancia de las variables β, a las q y a las p como funciones de t, de las α y de las constantes β, estas variables cumplirán con ecuaciones diferenciales que tienen forma hamiltoniana. Así una solución completa de la ecuación (6.2) permite obtener la solución general de las ecuaciones (6.16) y (6.17).

En resumen, una solución completa $G(q^1, \ldots, q^n, \alpha_1, \ldots, \alpha_n, t)$ de la ecuación (6.2), nos permite encontrar la solución general de las ecuaciones (6.16) y (6.17). Para hacer esto imponemos a las cantidades β definidas por (6.8) ser constantes, esto nos permite despejar las q y obtenerlas como funciones de t y de $2n$ constantes (las α y las β)

$$q^i = q^i(\alpha_1, \ldots, \alpha_n, \beta^1, \ldots, \beta^n, t); \tag{6.18}$$

substituyendo estas q en la definición (6.6) y obtenemos las p como función de las $2n$ constantes y el tiempo; estas q y p serán la solución general de las ecuaciones diferenciales ordinarias (6.16) y (6.17). Naturalmente se pueden relacionar las $2n$ constantes α y β con las condiciones iniciales de esas ecuaciones diferenciales.

Es importante señalar que para que todo esto tenga sentido la solución completa de la ecuación en derivadas parciales debe cumplir con la condición

$$\left| \frac{\partial^2 G}{\partial q^i \partial \alpha_j} \right| \neq 0, \tag{6.19}$$

de lo contrario no se podrían despejar las q de la ecuación (6.8), ni obtener la ecuación (6.16) a partir de (6.11). En general las ecuaciones en derivadas parciales de primer orden no necesariamente tiene soluciones completas que

cumplan con esta condición en todo el dominio de las variables q y α. En caso de no existir esas soluciones para la región de interés, no podremos encontrar las soluciones de (6.16) y (6.17) a partir de las de la ecuación en derivadas parciales.

Naturalmente, en caso de que la función K que define la ecuación en derivadas parciales sea la función de Hamilton de un sistema mecánico, la ecuación (6.2) será

$$H\left(q^1,\ldots,q^n,\frac{\partial G}{\partial q^1},\ldots,\frac{\partial G}{\partial q^n},t\right)=-\frac{\partial G}{\partial t}. \tag{6.20}$$

Esta ecuación, que puede ser obtenida de la hamiltoniana por una sustitución de los impulsos p por las derivadas parciales de G con respecto a las coordenadas q, se conoce como la ecuación de Hamilton Jacobi del sistema mecánico. Al aplicar el procedimiento general que hemos obtenido a este caso, obtendremos las trayectorias del sistema mecánico a partir del conocimento de una solución completa de esta ecuación.

Esto nos permite formular la mecánica de una manera distinta. Tomemos un sistema mecánico hamiltoniano, tendremos su variedad de configuración M, su espacio fase TM^* y su función hamiltoniana $H : TM^* \times R \to R$, donde hemos permitido a la hamiltoniana depender explícitamente del tiempo. En un sistema canónico de coordenadas (q,p) la hamiltoniana es

$$H(q^1,\ldots,q^n,p_1,\ldots,p_n,t). \tag{6.21}$$

A partir de esta hamiltoniana construyamos, de manera formal, una ecuación diferencial en derivadas parciales, la ecuación de Hamilton Jacobi,

$$H\left(q^1,\ldots,q^n,\frac{\partial G}{\partial q^1},\ldots,\frac{\partial G}{\partial q^n},t\right)=-\frac{\partial G}{\partial t}. \tag{6.22}$$

Una solución completa de esta ecuación que cumpla con la condición (6.19) nos proporciona la evolución temporal del sistema mecánico, vía el procedimiento de postular la constancia de las cantidades

$$\beta^i=\frac{\partial G}{\partial \alpha_i} \tag{6.23}$$

y la relación

$$p_i=\frac{\partial G}{\partial q^i}. \tag{6.24}$$

La argumentación que hemos seguido nos hace ver la equivalencia de esta formulación con la de Hamilton.

6.2. Ejemplos del proceso de Hamilton Jacobi.

Como un primer ejemplo estudiemos desde la perspectiva de Hamilton Jacobi a la partícula libre en dos dimensiones. La función hamiltoniana será

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2\right), \tag{6.25}$$

por tanto la ecuación de Hamilton Jacobi será

$$\frac{1}{2m}\left[\left(\frac{\partial G}{\partial x}\right)^2 + \left(\frac{\partial G}{\partial y}\right)^2\right] = -\frac{\partial G}{\partial t}. \tag{6.26}$$

Para encontrarle a esta ecuación una solución completa, esto es que contenga dos constantes arbitrarias no aditivas, seguiremos un proceso de separación de variables. Comencemos por proponer que la solución $G(x, y, t)$ pueda expresarse por $G(x, y, t) = W(x, y) + T(t)$. Al substituir esta expresión en la ecuación (6.26) encontramos que

$$\frac{1}{2m}\left[\left(\frac{\partial W(x,y)}{\partial x}\right)^2 + \left(\frac{\partial W(x,y)}{\partial y}\right)^2\right] = -\frac{\partial T(t)}{\partial t} \tag{6.27}$$

y como el lado izquierdo de la ecuación depende únicamente de las coordenadas espaciales y el derecho de la temporal ambos términos deben ser una misma constante a la que llamaremos α_1. Entonces

$$\frac{\partial T(t)}{\partial t} = -\alpha_1, \tag{6.28}$$

$$\frac{1}{2m}\left[\left(\frac{\partial W(x,y)}{\partial x}\right)^2 + \left(\frac{\partial W(x,y)}{\partial y}\right)^2\right] = \alpha_1. \tag{6.29}$$

De la primera de estas ecuaciones encontramos que

$$T(t) = -\alpha_1 t + c, \tag{6.30}$$

donde c es una constante de integración, que no incluiremos por ser en el contexto de la solución G una constante aditiva.

Para resolver la segunda ecuación supondremos que W puede separarse en una parte que depende sólo de x y otra que lo hace únicamente de y, esto es $W(x, y) = W_x(x) + W_y(y)$. Al substituir en la ecuación obtenemos

$$\frac{1}{2m}\left[\left(\frac{\partial W_x(x)}{\partial x}\right)^2 + \left(\frac{\partial W_y(y)}{\partial y}\right)^2\right] = \alpha_1, \tag{6.31}$$

esto implica, por la independencia de las variables x y y, que cada una de las dos derivadas deben ser constantes tales que cumplan con la ecuación, por tanto

$$\left(\frac{\partial W_x(x)}{\partial x}\right)^2 = \alpha_2 \tag{6.32}$$

y

$$\left(\frac{\partial W_y(y)}{\partial y}\right)^2 = -\alpha_2 + 2m\alpha_1. \tag{6.33}$$

Donde α_2 es una constante arbitraria. A partir de estas ecuaciones obtenemos que

$$\begin{aligned} W_x &= \sqrt{\alpha_2}x, \\ W_y &= \sqrt{-\alpha_2 + 2m\alpha_1}y, \end{aligned}$$

donde hemos omitido las constantes aditivas de integración.

Al sumar las tres contribuciones que hemos calculado obtenemos una solución completa de la ecuación de Hamilton Jacobi

$$G = \sqrt{\alpha_2}x + \sqrt{-\alpha_2 + 2m\alpha_1}y - \alpha_1 t. \tag{6.34}$$

A partir de esta solución completa de la ecuación de Hamilton Jacobi podemos encontrar la solución a las ecuaciones de Hamilton, esto es las ecuaciones de movimiento de la partícula libre, usando (6.8) y (6.6) e imponiendo que las β que aparecen en (6.8) sean constantes. De (6.8) tendremos

$$\frac{\partial G}{\partial \alpha_1} = \frac{my}{\sqrt{-\alpha_2 + 2m\alpha_1}} - t = \beta_1, \tag{6.35}$$

$$\frac{\partial G}{\partial \alpha_2} = \frac{x}{2\sqrt{\alpha_2}} - \frac{y}{2\sqrt{-\alpha_2 + 2m\alpha_1}} = \beta_2. \tag{6.36}$$

De la primera de estas ecuaciones encontramos que

$$y(t) = \frac{\sqrt{-\alpha_2 + 2m\alpha_1}}{m}(t + \beta_1) \tag{6.37}$$

y de la segunda que

$$x(t) = \frac{\sqrt{\alpha_2}}{m}t + \sqrt{\alpha_2}\left(\frac{\beta_1}{m} + 2\beta_2\right). \tag{6.38}$$

Los impulsos los encontramos a partir de (6.6), esto es

$$p_x = \frac{\partial G}{\partial x} = \sqrt{\alpha_2}, \tag{6.39}$$

$$p_y = \frac{\partial G}{\partial y} = \sqrt{-\alpha_2 + 2m\alpha_1}, \tag{6.40}$$

o sea que los dos impulsos son constantes y podemos ver que

$$\alpha_2 = p_x^2 \tag{6.41}$$

y que

$$\alpha_1 = \frac{1}{2m}\left(p_x^2 + p_y^2\right), \tag{6.42}$$

o sea la energía total.

Al substituir estas expresiones en las ecuaciones para x y y encontramos

$$x(t) = \frac{p_x}{m}t + p_x\left(\frac{\beta_1}{m} + 2\beta_2\right), \tag{6.43}$$

$$y(t) = \frac{p_y}{m}t + \frac{p_y}{m}\beta_1, \tag{6.44}$$

lo que nos permite identificar a β_1 con $y_0 m/p_y$ y a β_2 con $x_0/2p_x - y_0/2p_y$. Vemos claramente que las constantes α y β que obtenemos al aplicar el procedimiento de Hamilton Jacobi son una forma de expresar las condiciones iniciales del sistema mecánico en cuestión y que efectivamente, el conocer una solución completa de la ecuación de Hamilton Jacobi, nos permite, por un procedimiento puramente algebraico, encontrar la solución del sistema mecánico correspondiente.

Analicemos otro ejemplo, el problema de Kepler. En este sistema la hamiltoniana en coordenadas polares es

$$H = \frac{1}{2m}\left[p_r^2 + \frac{1}{r^2}p_\theta^2\right] - \frac{k}{r}, \tag{6.45}$$

por lo que la ecuación de Hamilton Jacobi será

$$\frac{1}{2m}\left[\left(\frac{\partial G}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial G}{\partial \theta}\right)^2\right] - \frac{k}{r} = -\frac{\partial G}{\partial t}. \tag{6.46}$$

Podemos ver que en esta ecuación el tiempo es también separable, por lo que proponemos una solución de la forma $G(r,\theta,t) = W(r,\theta) - \alpha_1 t$, al substituir en la ecuación de Hamilton Jacobi se ve que, para que G sea solución, W debe cumplir con

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2\right] - \frac{k}{r} = \alpha_1. \tag{6.47}$$

Esta ecuación es separable, por lo que proponemos una solución $W(r,\theta) = W_r(r) + W_\theta(\theta)$ que al ser substituida da

$$\frac{1}{2m}\left[\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2\right] - \frac{k}{r} = \alpha_1, \tag{6.48}$$

Toda la dependencia en θ de esta ecuación está en el término $\partial W_\theta/\partial\theta$, por lo que este debe ser constante, llamemos a esta constante α_2. Entonces

$$W_\theta(\theta) = \alpha_2\theta \tag{6.49}$$

más una constante aditiva de integración que no consideraremos por las razones ya expuestas. Finalmente W_r deberá, para que se cumpla la ecuación de Hamilton Jacobi, cumplir con

$$\frac{1}{2m}\left[\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{\alpha_2^2}{r^2}\right] - \frac{k}{r} = \alpha_1. \tag{6.50}$$

Para resolver esta ecuación despejamos el término de la derivada

$$\frac{\partial W_r}{\partial r} = \sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}} \tag{6.51}$$

e integramos

$$W_r = \int \sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}}dr. \tag{6.52}$$

La solución completa será

$$G = \int \sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}}dr + \alpha_2\theta - \alpha_1 t, \tag{6.53}$$

donde α_1 y α_2 son las dos constantes arbitrarias no aditivas que se requieren.

La integral puede hacerse y expresarse analíticamente, pero no es necesario pues únicamente se utilizan las derivadas de G. De (6.8) tendremos

$$\beta_1 = \frac{\partial G}{\partial \alpha_1} = \int \frac{m}{\sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}}}dr - t, \tag{6.54}$$

$$\beta_2 = \frac{\partial G}{\partial \alpha_2} = -\int \frac{\alpha_2}{r^2\sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}}}dr + \theta. \tag{6.55}$$

De la primera de estas ecuaciones, considerando a β_1 como constante, despejamos r, que nos quedará en función del tiempo y tres constantes arbitrarias las dos α y la β_1. En la segunda ecuación substituimos la $r(t)$ que hayamos obtenido de la primera y despejamos θ, que quedará en función del tiempo y de las cuatro constantes. Notamos que esta segunda ecuación es la ecuación de la órbita, pues relaciona a la r con la θ y que la integral que aparece es la misma que obtuvimos en el capítulo tres al integrar directamente las ecuaciones de movimiento.

Utilizamos ahora (6.6)

$$p_r = \frac{\partial G}{\partial r} = \sqrt{2m\alpha_1 + 2m\frac{k}{r} - \frac{\alpha_2^2}{r^2}}, \tag{6.56}$$

$$p_\theta = \frac{\partial G}{\partial \theta} = \alpha_2, \tag{6.57}$$

al substituir $r(t)$ en la primera de estas ecuaciones obtenemos p_r como función del tiempo y de las cuatro constantes; la segunda ecuación nos indica que p_θ es constante (conservación del impulso angular) e igual a α_2. Es fácil hacer ver que α_1 es la energía total.

Con esto tendremos resuelto totalmente el problema de Kepler por medio del método de Hamilton Jacobi. Remarcamos que únicamente hemos resuelto una ecuación diferencia parcial para obtener G y que a partir de esta solución, por un proceso algebraico, obtenemos la solución completa del problema mecánico. Las constantes que aparecen, las α y las β, no son otra cosa que una manera de expresar las condiciones iniciales que siempre están presentes en la solución de un sistema mecánico.

Notemos que, en el caso general, cuando la hamiltoniana es independiente del tiempo, siempre podemos separar el tiempo y tener soluciones de la forma

$$G(q_1, \ldots, q_n, t, \alpha_1, \ldots \alpha_n) = W(q_1, \ldots, q_n, \alpha_1, \ldots \alpha_n) - \alpha_1 t, \tag{6.58}$$

naturalmente también podremos tener muchas otras soluciones en las que el tiempo no queda separado.

Podemos ver a la ecuación de Hamilton Jacobi como una nueva formulación de la mecánica matemáticamente equivalente a la formulación hamiltoniana. Desde esta perspectiva una función del espacio de configuración, el tiempo y n constantes arbitrarias o parámetros no aditivos, resulta ser el elemento definitorio, pues contiene totalmente a la evolución temporal del sistema. La "ley básica del movimiento", que representa el papel que en otras formulaciones tienen las ecuaciones de Newton, las de Lagrange o las de

Hamilton, lo tiene ahora la ecuación de Hamilton Jacobi, cuyas soluciones proporcionan la función que determina la evolución temporal del sistema.

Es de destacar que esta formulación presenta dos diferencias fundamentales con las otras: se basa en una ecuación en derivadas parciales y no en ecuaciones diferenciales ordinarias y se introduce un elemento nuevo, que no tiene contraparte en las demás formulaciones, una función, solución completa de esa ecuación parcial. De estas dos diferencias, que en realidad están muy relacionadas la una con la otra, es la segunda la que llama más la atención, puesto que nos lleva a preguntar sobre el significado o contenido físico de la función G, contenido o interpretación que no puede tener cabida en las otras formulaciones, ya que en ellas no aparece nada similar a ella. Pasemos pues a ver que sentido físico podemos darle a esta función.

Pensemos en un caso sencillo, el de una partícula de masa m que se encuentra en el espacio físico ordinario, espacio al que le asociamos una métrica euclidiana y lo describimos haciendo uso de un sistema cartesiano (x, y, z). La partícula esta en presencia de un potencial que es función únicamente del espacio físico. Notemos que en el espacio físico y en las coordenadas que estamos usando el tensor métrico es δ_{ij}, por eso dada una función $g(x, y, z)$ las componentes, en las bases inducidas por el sistema de coordenadas, de su diferencial y de su gradiente son las mismas cantidades,

$$\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right). \tag{6.59}$$

Por otro lado el tensor de inercia es $m\delta_{ij}$ por lo que la relación entre las componentes de la velocidad $(\dot{x}, \dot{y}, \dot{z})$ y las del impulso (p_x, p_y, p_z) es simplemente

$$\begin{aligned} p_x &= m\dot{x}, \\ p_y &= m\dot{y}, \\ p_z &= m\dot{z}. \end{aligned}$$

La hamiltoniana de este sistema es independiente del tiempo por lo que podemos separarlo y utilizar una solución completa de la ecuación de Hamilton Jacobi de la forma

$$G(x, y, z, t) = W(x, y, z) - \alpha_1 t, \tag{6.60}$$

donde W dependerá de tres constantes no aditivas α_1, α_2 y α_3.

De la ecuación (6.8) tenemos que al considerar las β como constantes las ecuaciones

$$\beta_1 = \frac{\partial W}{\partial \alpha_1} - t, \tag{6.61}$$

$$\beta_2 = \frac{\partial W}{\partial \alpha_2}, \tag{6.62}$$

$$\beta_3 = \frac{\partial W}{\partial \alpha_3}, \tag{6.63}$$

nos proporcionan las curvas que traza la partícula en el espacio físico en su movimiento. Obtenemos una de estas trayectorias para cada conjunto de valores de las α y de las β, que únicamente reflejan las condiciones iniciales. Los vectores tangentes a estas curvas tendrán como componentes a $\dot{x}$, $\dot{y}$ y $\dot{z}$.

Por otro lado el impulso, que en este caso es un vector paralelo al vector tangente a las trayectorias, está dado por (6.6), esto es

$$p_x = \frac{\partial W}{\partial x}, \tag{6.64}$$

$$p_y = \frac{\partial W}{\partial y}, \tag{6.65}$$

$$p_z = \frac{\partial W}{\partial z}, \tag{6.66}$$

lo que, en este caso, podemos identificar con las componentes del gradiente de W. Esto nos hace ver que las trayectorias son un conjunto de curvas tangentes al campo vectorial gradiente de W y por tanto son curvas perpendiculares a las superficies de W constante, como se muestra en la figura 6.1 para el caso de una partícula libre. Esto es totalmente análogo al caso en que tenemos un campo escalar y lo vemos por medio de sus lineas de flujo, que no son otra cosa que curvas perpendiculares a las superficies de campo constante.

Así podemos ver a W como un campo escalar en la variedad de configuración parametrizado por las alfas, y a las trayectorias como lineas de flujo de ese campo; podemos, no sin hacer ciertos abusos del lenguaje, decir que el sistema mecánico genera un campo escalar cuya ecuación es la de Hamilton Jacobi y que la partícula sigue, en su movimiento, las lineas de fuerza de ese campo, trazando así las trayectorias, todo ello en la variedad de configuración.

En el caso de una partícula libre en dos dimensiones que hemos estudiado tenemos que una solución completa de la ecuación de Hamilton Jacobi es

$$G = \sqrt{\alpha_2}x + \sqrt{-\alpha_2 + 2m\alpha_1}y - \alpha_1 t \tag{6.67}$$

y por tanto las superficies de W constante serán

$$\sqrt{\alpha_2}x + \sqrt{-\alpha_2 + 2m\alpha_1}y = c \tag{6.68}$$

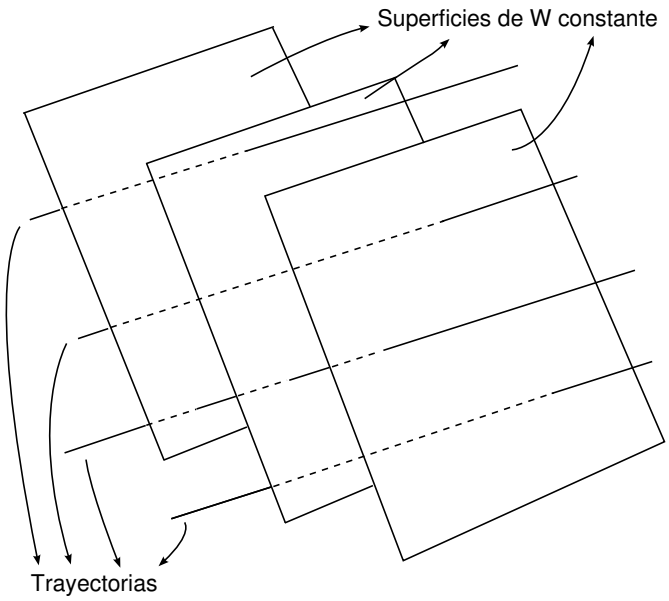


Figura 6.1: La solución, obtenida por separación de variables en coordenadas cartesianas, de la ecuación de Hamilton Jacobi de la partícula libre, vista como un campo escalar. Los planos son las superficies de W constante, la orientación de estos planos queda determinada por los valores de las constantes α, las trayectorias forman un haz de rectas perpendiculares a los planos. Las constantes β determinan una trayectoria específica entre todas las que conforman el haz.

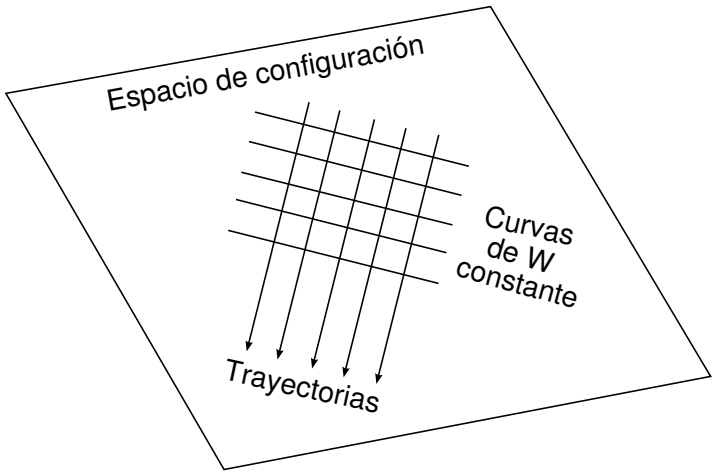


Figura 6.2: Las trayectorias y las curves de W constante para una partícula libre que se mueve en un plano.

donde hemos llamado c a la constante. Estas superficies, en este caso de dos grados de libertad, serán rectas paralelas de pendiente

$$-\frac{1}{\sqrt{-1+2m\frac{\alpha_1}{\alpha_2}}}. \tag{6.69}$$

Las trayectorias del sistema dadas por (6.38) y (6.37) son rectas de pendiente

$$\sqrt{-1+2m\frac{\alpha_1}{\alpha_2}}, \tag{6.70}$$

esto es, rectas paralelas perpendiculares a las anteriores en una métrica euclidiana correspondiente a que, en este caso, el tensor de inercia es $m\delta_{ij}$.

Tomemos un segundo ejemplo, el tiro parabólico. La hamiltoniana en coordenadas cartesianas es

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2\right) + mgy, \tag{6.71}$$

por lo que la ecuación de Hamilton Jacobi será

$$\frac{1}{2m}\left[\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2\right] + mgy = -\frac{\partial F}{\partial t}, \tag{6.72}$$

al separar el tiempo y proponer que $F = W(x,y) - \alpha_1 t$ tendremos que

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2\right] + mgy = \alpha_1. \tag{6.73}$$

Esta última ecuación es separable al proponer que $W(x,y) = W_x(x) + W_y(y)$ y obtenemos que se deben cumplir las ecuaciones

$$\frac{\partial W_x}{\partial x} = \alpha_2 \tag{6.74}$$

y

$$\frac{1}{2m}\left[\alpha_2^2 + \left(\frac{\partial W_y}{\partial y}\right)^2\right] + mgy = \alpha_1. \tag{6.75}$$

De la primera

$$W_x = \alpha_2 x \tag{6.76}$$

y de la segunda

$$W_y = \mp\frac{1}{3m^2 g}\left[2m\alpha_1 - \alpha_2^2 - 2m^2 gy\right]^{\frac{3}{2}}, \tag{6.77}$$

donde hemos omitido las constantes aditivas. Finalmente tendremos que

$$W = \alpha_2 x \mp \frac{1}{3m^2 g}\left[2m\alpha_1 - \alpha_2^2 - 2m^2 gy\right]^{\frac{3}{2}}. \tag{6.78}$$

A partir de esta última expresión podemos seguir el procedimiento de Hamilton Jacobi para encontrar que

$$x = \frac{\alpha_2}{m}(t + \beta_1) + \beta_2 \tag{6.79}$$

y

$$y = -\frac{g}{2}(t + \beta_1)^2 + \frac{2m\alpha_1 - \alpha_2^2}{2m^2 g}, \tag{6.80}$$

que es la trayectoria de un tiro parabólico en el espacio de configuración.

Ahora podemos trazar, usando (6.78), las curvas de W constante o curvas de nivel del campo escalar W y las trayectorias, todo ello en el espacio de configuración, esto es, en el plano (x,y); en la figura 6.3 se muestran estas gráficas. La expresión (6.78) muestra que W tiene dos ramas, correspondientes a los signos positivo o negativo del radical, por tanto las curvas $W = cst.$ también tendrán dos ramas

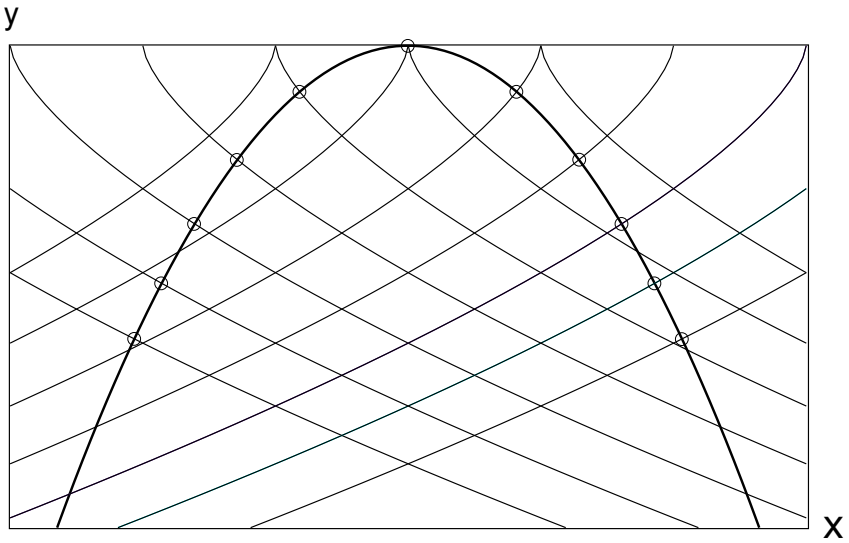


Figura 6.3: Una trayectoria de un tiro parabólico trazada sobre el fondo de las curvas de nivel del campo escalar W dado por (6.78). La trayectoria se ha dibujado con una linea gruesa y únicamente se han trazado las dos ramas de las curvas de nivel para tres casos. La trayectoria sigue unas ramas cuando asciende y las otras cuando desciende, por eso las dos ramas de un mismo nivel generan una cúspide en la altura máxima que alcanzan las trayectorias. Con unos pequeños círculos se han enmarcado los puntos en los que se muestra la perpendicularidad de la trayectoria con las curvas de nivel de W.

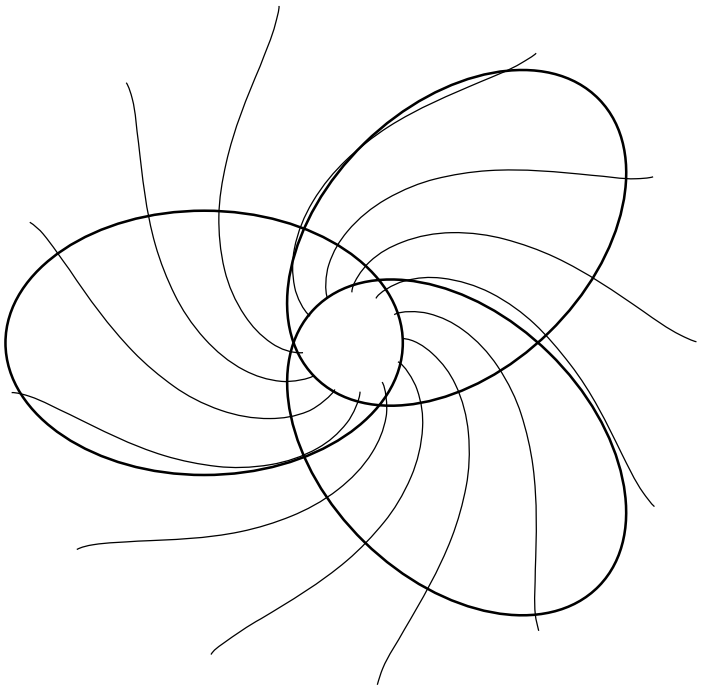


Figura 6.4: Curvas de nivel del campo escalar W para el caso planetario, calculadas por separación de variables en coordenadas polares. En traza gruesa se tienen tres órbitas elípticas. En traza delgada las curvas de nivel, de éstas únicamente se ha trazado la rama ascendente de W_r.

En la figura 6.4 hemos trazado las curvas de nivel del campo escalar W para el movimiento planetario. W es el campo dado por $W = W_\theta + W_r$, donde w_θ y W_r salen de (6.49) y (6.52). Para trazar esta figura se trasladan las coordenadas polares a cartesianas, pero se debe tener en consideración que la separación de variables se hizo en coordenadas polares. Para no complicar la figura únicamente se han trazado las ramas correspondientes al signo positivo del radical en (6.52), estas ramas sólo generarían las ramas ascendentes (r en crecimiento) de las órbitas, se han trazado tres órbitas elípticas completas.

En un sistema en el que la hamiltoniana dependa expresamente del tiempo o cuando tenemos soluciones en las que no se ha separado el tiempo, tendremos que usar el espacio extendido y contemplar las órbitas desplegadas en el tiempo como curvas transversales a las hipersuperficies dadas por $G = \text{cst.}$; en general no tendrá sentido la función W que depende únicamente de las coordenadas.

En el caso de un sistema mecánico general en el que la hamiltoniana es independiente del tiempo y en que tenemos una solución de la forma (6.60), tendremos una variedad de configuración en lugar del espacio físico tridimensional y la energía cinética será de la forma ref****, esto es, en vez de la masa de la partícula tendremos un tensor de inercia. La ecuación de Hamilton Jacobi (6.2) se expresa en una carta, o sea será una ecuación local, las trayectorias las encontramos utilizando (6.6) y los impulsos estarán determinados por (6.8), pero ahora la relación entre estos impulsos, que debemos ver como las componentes de la diferencial de W, y las componentes de los vectores tangentes a las trayectorias, estará dada por el tensor de inercia y, para hablar de la perpendicularidad de las trayectorias a las hipersuperficies de $W = \text{cst.}$ debemos usar a ese tensor de inercia como tensor métrico en la variedad de configuración.[1] Veamos esto con más detenimiento.

En el caso general independiente del tiempo tenemos una variedad de configuración M y una función lagrangiana $L = T - V : TM \to R$ independiente del tiempo. Definimos en M un producto interno utilizando al tensor de energía cinética I_{ij} como tensor métrico; la energía cinética será entonces

$$T = \frac{1}{2} < v, v >, \tag{6.81}$$

donde hemos denotado con v al campo vectorial tangente a las trayectorias del sistema en la variedad M, esto es, curvas que cumplen con las ecuaciones de Lagrange. Nos preguntamos si existe algún campo escalar $W : M \to R$ tal que su gradiente sea un campo vectorial cuyo flujo en M esté compuesto por soluciones del sistema mecánico y de ser así como determinarlo.

Recordamos que para cualquier campo vectorial X y cualquier función F

$$dF[X] = X[F] \tag{6.82}$$

y que la aplicación de un campo vectorial a una función es la derivada direccional de la función en la dirección del campo. Tendremos entonces que

$$dW[\text{grad}W] = \text{grad}W[W], \tag{6.83}$$

[1] Recordamos que mientras no definamos un tensor métrico en la variedad de configuración, no tiene sentido hablar de perpendicularidad, únicamente podríamos usar la noción de transversalidad. Asimismo, si definimos como tensor métrico de la variedad de configuración a uno que no sea el de inercia del sistema, la relación entre las p y las $\dot{q}$ no será la de una forma diferencial y un campo vectorial en el sentido de la métrica, por lo que no tendríamos esa perpendicularidad entre las trayectorias y las hipersuperficies de $W = \text{cst.}$, aunque si se mantendría la noción de transversalidad.

ahora queremos que grad$W$ sea un campo tangente a soluciones del sistema, por lo que

$$\text{grad}W[W] = \frac{d}{dt}W \tag{6.84}$$

y por tanto

$$dW[\text{grad}W] = \frac{d}{dt}W; \tag{6.85}$$

vemos también que la energía cinética (6.81) será

$$T = \frac{1}{2}dW[\text{grad}W]. \tag{6.86}$$

Por otro lado la diferencial de W será la forma diferencial que se corresponde con el gradiente de W (que es el campo de 'velocidades') según una métrica dada por el tensor de inercia, esto es, $dW =< \bullet, \text{grad}W >$ es la forma diferencial de los impulsos, por lo que las ecuaciones de Lagrange tendrán la forma

$$\frac{d}{dt}[dW] - dL = 0; \tag{6.87}$$

como el tiempo es sólo un parámetro podemos escribir

$$d\left(\frac{d}{dt}W\right) = dL, \tag{6.88}$$

e integrando

$$\frac{d}{dt}W = L + c, \tag{6.89}$$

donde c es una constante de integración arbitraria.

Ahora de (6.85) y del hecho de que

$$L = \frac{1}{2}dW[\text{grad}W] - V, \tag{6.90}$$

tenemos que, para que se cumplan las ecuaciones de Lagrange, W debe cumplir con

$$dW[\text{grad}W] = \frac{1}{2}dW[\text{grad}W] - V + c, \tag{6.91}$$

o lo que es lo mismo

$$\frac{1}{2}dW[\text{grad}W] + V = c, \tag{6.92}$$

que no es otra cosa que la ecuación de Hamilton Jacobi cuando se ha separado el tiempo, c sería la constante arbitraría que en el proceso de separación del tiempo hemos llamado α_1.

Si la lagrangiana depende del tiempo o usamos una solución en la que el tiempo no se ha separado, tendríamos que utilizar una variedad de configuración extendida con una dimensión temporal y hacer las generalizaciones que se hicieron en el ejemplo de una partícula. Veamos esto utilizando coordenadas para el caso en que la energía cinética se expresa como

$$T = \frac{1}{2}\sum_{ij}\left(I_{ij}(q)\dot{q}^i\dot{q}^j\right). \tag{6.93}$$

La lagrangiana será

$$L = \frac{1}{2}\sum_{ij}\left(I_{ij}(q)\dot{q}^i\dot{q}^j\right) - V(q,t) \tag{6.94}$$

y las ecuaciones de Lagrange

$$\frac{d}{dt}\left(\sum_j I_{kj}\dot{q}^j\right) - \frac{1}{2}\sum_{ij}\left(\frac{\partial I_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j\right) + \frac{\partial V}{\partial q^k} = 0. \tag{6.95}$$

Estamos buscando un campo vectorial de gradiente, $v = \text{grad}F(q,t)$, tal que las ecuaciones $\dot{q}^i = v^i$ tengan soluciones $q^i(t)$ que sean también soluciones de las ecuaciones de Lagrange (6.95), naturalmente para hablar de campo de gradiente debemos tener una métrica en la variedad de configuración, tomemos al tensor de inercia como tal, entonces

$$\dot{q}^i = v^i = \sum_j I^{ij}\frac{\partial F}{\partial q^j} \tag{6.96}$$

o en sentido inverso

$$\frac{\partial F}{\partial q^k} = \sum_j I_{kj}\dot{q}^j. \tag{6.97}$$

Para que las ecuaciones (6.96) tengan soluciones que satisfagan las ecuaciones de Lagrange, deben satisfacer (6.95), por lo que al tomar en cuenta (6.97) tendremos

$$\frac{d}{dt}\left(\frac{\partial F}{\partial q^k}\right) - \frac{1}{2}\sum_{ij}\left(\frac{\partial I_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j\right) + \frac{\partial V}{\partial q^k} = 0, \tag{6.98}$$

pero de (6.97)

$$\sum_j \frac{\partial I_{ij}}{\partial q^k}\dot{q}^j = \frac{\partial^2 F}{\partial q^k \partial q^i}, \tag{6.99}$$

por lo que F debe cumplir con

$$\frac{d}{dt}\left(\frac{\partial F}{\partial q^k}\right) - \frac{1}{2}\sum_i \left(\frac{\partial^2 F}{\partial q^i \partial q^k}\dot{q}^i\right) + \frac{\partial V}{\partial q^k} = 0. \tag{6.100}$$

Por otro lado

$$\frac{d}{dt}\left(\frac{\partial F}{\partial q^k}\right) = \sum_i \frac{\partial^2 F}{\partial q^i \partial q^k}\dot{q}^i + \frac{\partial^2 F}{\partial t \partial q^k}, \tag{6.101}$$

por lo que al substituir en (6.100) obtenemos

$$\frac{1}{2}\sum_i \left(\frac{\partial^2 F}{\partial q^i \partial q^k}\dot{q}^i\right) + \frac{\partial^2 F}{\partial t \partial q^k} + \frac{\partial V}{\partial q^k} = 0 \tag{6.102}$$

y usando (6.96) tendremos

$$\frac{\partial}{\partial q^k}\left[\frac{1}{2}\sum_{ij} I^{ij}\frac{\partial F}{\partial q^i}\frac{\partial F}{\partial q^j} + \frac{\partial F}{\partial t} + V\right] = 0, \tag{6.103}$$

o, por tener que ser esto valido para todos los valores de k,

$$\frac{1}{2}\sum_{ij} I^{ij}\frac{\partial F}{\partial q^i}\frac{\partial F}{\partial q^j} + V + c = -\frac{\partial F}{\partial t}; \tag{6.104}$$

que, al hacer la constante aditiva arbitraria cero, es la ecuación de Hamilton Jacobi.

Hemos así dado una posible interpretación a las soluciones de la ecuación de Hamilton Jacobi, podemos contemplarlas como campos escalares en la variedad de configuración cuyo campo de gradiente, en la métrica establecida por el tensor de inercia, tiene como lineas de flujo soluciones del sistema mecánico. Naturalmente una solución de la ecuación de Hamilton Jacobi no puede, por esta vía, generar todas las soluciones del sistema mecánico, ya que únicamente tendríamos n condiciones iniciales a disposición; tendríamos que tener una solución independiente por cada conjunto de n valores arbitrarios que cubrieran las otras condiciones iniciales, pero esto es lo que proporciona una solución completa, por lo que podemos contemplar a una solución completa como un campo escalar parametrizado por las constantes no aditivas que la conforman.

6.3. La aproximación de la eikonal y la ecuación de Hamilton Jacobi.

En el capitulo *** encontramos una analogía entre la mecánica y la óptica geométrica: ambas pueden ser formuladas por medio de un principio variacional, el de Hamilton para la mecánica y el de Fermat para la óptica. Esta analogía puede ser extendida al contemplarlas desde la formulación de Hamilton Jacobi una y desde un proceso de aproximación que nos lleva, de la visión ondulatoria de la luz, a la óptica geométrica, proceso que da lugar a la llamada ecuación de la eikonal. Veamos esto con detalle.

La óptica geométrica estudia las curvas que trazan en el espacio los rayos de luz cuando viajan en un medio óptico. Sabemos que esto es sólo una aproximación al fenómeno óptico real, ya que la luz son ondas electromagnéticas que interaccionan con el medio óptico, al considerar a la luz como rayos, se dejan a un lado todos los fenómenos que tienen como origen la naturaleza ondulatoria, como son los de difracción, interferencia o polarización. La óptica geométrica es una aproximación valida únicamente cuando las distancias características en que cambian las propiedades ópticas del medio son grandes comparadas con la longitud de onda de la luz, por lo que se trata de una aproximación de longitud de onda pequeña, esto es, describe el comportamiento de la luz cuando su longitud de onda tiende a cero. Al tomar las ecuaciones de onda y encontrar, a partir de ellas, una aproximación para longitud de onda pequeña, se obtiene la llamada ecuación o aproximación de la eikonal.

Para el objeto de este curso estudiaremos un caso simple, el de una onda escalar, en un medio heterogéneo pero isotrópico e invariante en el tiempo, esto es una onda en la que la magnitud de su velocidad depende del medio, pero no de la dirección de propagación, de la longitud de onda, del tiempo o de otros factores.

La ecuación de onda en un sistema cartesiano de coordenadas será

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}-\frac{n^2(x,y,z)}{c^2}\frac{\partial^2\phi}{\partial t^2}=0, \tag{6.105}$$

donde $n(x,y,z)$ es el índice de refracción, c la velocidad en el vacío y $\phi(x,y,z,t)$ puede verse como una de las componentes del campo eléctrico o magnético.[2]

Cuando tenemos un medio homogéneo, esto es, uno en el que n es cons-

[2]Esto si nos mantenemos en el mismo sistema de coordenadas cartesianas, ya que ϕ debería representar un escalar y las componentes de los campos vectoriales no son escalares.

tante, podemos tener soluciones llamadas de onda plana

$$\phi = Ae^{i\frac{2\pi}{\epsilon}(\overline{k}\cdot\overline{x}-ct)} \tag{6.106}$$

donde A es una amplitud constante, $\overline{k}$ es un vector de onda constante, $\overline{x}$ el vector de posición y ϵ un parámetro de escala auxiliar. Al substituir en la ecuación de onda encontramos que, para que esta se cumpla, debemos tener

$$\left|\overline{k}\right|^2 = n^2. \tag{6.107}$$

A estas soluciones se les llama de onda plana porque las superficies de fase constante o frentes de onda, es decir aquellas sobre las que el argumento de la exponencial es constante

$$\frac{2\pi}{\epsilon}\left(\overline{k}\cdot\overline{x}-ct\right) = \text{cst.}, \tag{6.108}$$

son planos perpendiculares al vector de onda que, con el tiempo, se trasladan paralelos a si mismos con la velocidad de fase. La longitud de onda λ es la distancia entre dos de estos planos que difieran por una fase 2π a un tiempo fijo

$$\lambda = \frac{\epsilon}{\left|\overline{k}\right|} = \frac{\epsilon}{n}. \tag{6.109}$$

El período τ será el tiempo en el que, en un punto fijo del espacio, la fase cambia por 2π, está dado por

$$\tau = \frac{\epsilon}{c} \tag{6.110}$$

y la frecuencia angular será entonces

$$\omega = \frac{2\pi}{\tau} = \frac{2\pi c}{\epsilon}. \tag{6.111}$$

Estas soluciones son de gran utilidad dado que el análisis de Fourier nos permite estudiar muchas otras soluciones con base en estas ondas planas.

Cuando el medio no es homogéneo, esto es cuando el índice de refracción n es función de la posición, las ondas planas ya no son solución de la ecuación de ondas. La inhomogeneidad hace que cualquier frente de onda plano se distorsione al avanzar con el tiempo y deja de ser plano.

Por otra parte es muy conveniente el poder analizar soluciones que localmente puedan verse como ondas planas, es decir, que en una vecindad de cada punto del espacio se puedan aproximar como una onda plana y se pueda

definir un vector de onda, que naturalmente dependerá del punto en cuestión, y una frecuencia que supondremos no varía con la posición. Proponemos entonces soluciones de la forma

$$\phi = A(\overline{x})e^{i\frac{2\pi}{\epsilon}(S(\overline{x})-ct)} \tag{6.112}$$

en donde hemos supuesto una amplitud variable y compleja y hemos mantenido el concepto de fase de la onda. En esta ocasión la fase varía con una función general de la posición $S(\overline{x})$, aunque hemos conservado la forma de la dependencia temporal.

Al substituir está propuesta de solución en la ecuación de ondas 6.105 encontramos las ecuaciones que deben cumplir A y S para que ϕ sea solución. Hagamos este cálculo.

El laplaciano es la divergencia del gradiente, calculemos pues el gradiente de ϕ

$$\bigtriangledown\phi = \left[\bigtriangledown A + i\frac{2\pi}{\epsilon}A \bigtriangledown S\right] e^{i\frac{2\pi}{\epsilon}(S(\overline{x})-ct)}. \tag{6.113}$$

Calculemos ahora la divergencia de esta expresión

$$\bigtriangledown^2\phi = \bigtriangledown\cdot\bigtriangledown\phi = \left[\bigtriangledown^2 A - \frac{4\pi^2}{\epsilon^2}A \bigtriangledown S\cdot\bigtriangledown S + i\frac{2\pi}{\epsilon}\left(2\bigtriangledown A\cdot\bigtriangledown S + A\bigtriangledown^2 S\right)\right] e^{i\frac{2\pi}{\epsilon}(S(\overline{x})-ct)}. \tag{6.114}$$

Por otra parte tendremos que

$$\frac{\partial^2\phi}{\partial t^2} = -\frac{4\pi^2}{\epsilon^2}Ac^2 e^{i\frac{2\pi}{\epsilon}(S(\overline{x})-ct)}, \tag{6.115}$$

por lo que, para que se cumpla la ecuación de ondas, debemos tener que

$$\epsilon^2\bigtriangledown^2 A - 4\pi^2 A\bigtriangledown S\cdot\bigtriangledown S + 4\pi^2 An^2 + i2\pi\epsilon\left(2\bigtriangledown A\cdot\bigtriangledown S + A\bigtriangledown^2 S\right) = 0. \tag{6.116}$$

Igualando a cero tanto la parte real como la imaginaria obtenemos dos ecuaciones que nos proporcionan $A(\overline{x})$ y $S(\overline{x})$.

Una vez que tenemos una solución de la forma (6.112) podemos recuperar los conceptos de longitud de onda y frecuencia angular. Para ello nos fijamos en un punto $\overline{x}_0$ del espacio, en él la fase tendrá un cierto valor f_0 y pasará una superficie de fase constante o frente de onda dada por

$$\frac{2\pi}{\epsilon}(S(\overline{x}) - ct) = f_0 = \frac{2\pi}{\epsilon}(S(\overline{x}_0) - ct), \tag{6.117}$$

que en general ya no será un plano, pero que se irá trasladando con el tiempo. Por otra parte habrá otras superficies en que la fase, a un tiempo fijo, difiera

de ésta por 2π, la distancia entre estas superficies contendrá el concepto de longitud de onda, pero en esta ocasión esta distancia no es constante, por lo que definiremos la longitud de onda como la distancia desde el punto $\overline{x}_0$ en el primer frente de onda al segundo frente onda, medida a lo largo de la linea perpendicular al primero en el punto $\overline{x}_0$. La perpendicular al frente de onda en el punto x_0 está determinada por el gradiente de S valuado en ese punto, denotaremos a este gradiente por por $\bigtriangledown S_0$. La longitud de onda en el punto $\overline{x}_0$ estará determinada por la ecuación

$$\frac{2\pi}{\epsilon}\left[S\left(\overline{x}_0+\frac{\bigtriangledown S_0}{|\bigtriangledown S_0|}\lambda\right)-ct\right]-\frac{2\pi}{\epsilon}\left(s(\overline{x}_0)-ct\right)=2\pi; \tag{6.118}$$

desarrollando en serie S en el primer término y quedándonos a primer orden tendremos

$$\frac{2\pi}{\epsilon}\left|\bigtriangledown S_0\right|\lambda=2\pi, \tag{6.119}$$

por lo que

$$\lambda=\frac{\epsilon}{|\bigtriangledown S_0|}. \tag{6.120}$$

Notamos que ahora la longitud de onda es dependiente del punto del espacio en que estemos, pues podemos contemplar al gradiente de S como un vector de onda que depende del punto, por otro lado el parámetro ϵ nos es útil como una escala que relaciona la longitud de onda con el cambio de S en el espacio, esto es con la magnitud de su gradiente, por lo que la aproximación de longitud de onda pequeña se puede tomar al hacer tender ϵ a cero. El período, la frecuencia angular y la velocidad de fase estarán dadas por

$$\tau=\frac{\epsilon}{c}, \tag{6.121}$$

$$\omega=\frac{2\pi c}{\epsilon} \tag{6.122}$$

$$v_f=\frac{c}{|\bigtriangledown S_0|}, \tag{6.123}$$

vemos, como era de esperarse, que el período y la frecuencia son constantes, mientras que la velocidad de fase es dependiente del espacio.

Ahora podemos retomar la ecuación (6.116) para encontrar la ecuación para S en la aproximación de longitud de onda pequeña, esto es cuando ϵ tiende a cero y son despreciables los términos que la contienen, obtenemos

$$\bigtriangledown S\cdot\bigtriangledown S=n^2, \tag{6.124}$$

la que en coordenadas cartesianas es

$$\left(\frac{\partial S_x}{\partial x}\right)^2 + \left(\frac{\partial S_y}{\partial y}\right)^2 + \left(\frac{\partial S_z}{\partial z}\right)^2 = n^2. \tag{6.125}$$

Esta ecuación, en la óptica, se conoce como ecuación de la eikonal. Nos proporciona S en la aproximación de longitud de onda pequeña o aproximación de la eikonal. Las superficies de S constante serán las que nos proporcionan la geometría de la transmisión de la luz cuando no se toman en cuenta, por razón de la aproximación que hemos hecho, los efectos ondulatorios. Esta geometría la podemos observar como una propagación a lo largo de las curvas perpendiculares a las superficies de S constante, lo que se conoce como congruencia de curvas perpendiculares a la familia de superficies de S constante. En el caso de la óptica estas curvas son los rayos y en cada punto el vector tangente a estos rayos será el gradiente de S valuado en el punto.

Notamos de inmediato la gran semejanza que hay entre la ecuación de la eikonal y la de Hamilton Jacobi y entre las trayectorias de un sistema mecánico, que son curvas transversales a las superficies en que las soluciones de la ecuación de Hamilton Jacobi se hacen constantes y los rayos de luz que son curvas perpendiculares a las superficies de S o fase constante, cuando S es solución de la ecuación de la eikonal.

Podemos directamente calcular la ecuación de la curva que trazan estos rayos. Para ello consideramos el vector de posición $\overline{r}$ y la ecuación de dicha curva en forma paramétrica utilizando un parámetro que llamaremos s, esto es la curva $\overline{r}(s)$. Esta curva debe ser perpendicular a las superficies de S constante, esto es, el gradiente de S debe ser tangente a esa curva en todos sus puntos, por lo que la curva será solución de la ecuación diferencial

$$\frac{d\overline{r}}{ds} = \bigtriangledown S, \tag{6.126}$$

o en componentes

$$\begin{aligned} \frac{dx}{ds} &= \frac{\partial S}{\partial x}, \\ \frac{dy}{ds} &= \frac{\partial S}{\partial y}, \\ \frac{dz}{ds} &= \frac{\partial S}{\partial z}. \end{aligned} \tag{6.127}$$

Esto implica, al tomar en cuenta la ecuación de la eikonal, que

$$\left|\frac{d\overline{r}}{ds}\right|^2 = \bigtriangledown S \cdot \bigtriangledown S = n^2. \tag{6.128}$$

Ahora si tomamos el operador

$$\frac{d}{ds} \tag{6.129}$$

tendremos que

$$\frac{d}{ds} = \frac{d\overline{r}}{ds} \cdot \bigtriangledown, \tag{6.130}$$

aplicando esto a una de las componentes de $d\overline{r}/ds$, digamos a la componente x, y tomando en consideración (6.127) obtenemos

$$\begin{aligned} \frac{d}{ds}\left(\frac{dx}{ds}\right) &= \frac{d}{ds}\left(\frac{\partial S}{\partial x}\right) = \frac{d\overline{r}}{ds} \cdot \bigtriangledown \left(\frac{\partial S}{\partial x}\right) = \bigtriangledown S \cdot \bigtriangledown \left(\frac{\partial S}{\partial x}\right) \\ &= \frac{\partial S}{\partial x}\frac{\partial^2 S}{\partial x^2} + \frac{\partial S}{\partial y}\frac{\partial^2 S}{\partial y \partial x} + \frac{\partial S}{\partial z}\frac{\partial^2 S}{\partial z \partial x} \\ &= \frac{1}{2}\frac{\partial}{\partial x}\left(\bigtriangledown S \cdot \bigtriangledown S\right) = \frac{1}{2}\frac{\partial n^2}{\partial x}, \end{aligned}$$

en donde hemos hecho uso de la ecuación de la eikonal para obtener la última igualdad. Como sucede lo mismo para las otras componentes tendremos que

$$\frac{d^2\overline{r}}{ds^2} = \frac{1}{2} \bigtriangledown n^2, \tag{6.131}$$

ecuación diferencial que en compañía de la condición (6.128) nos proporciona las curvas que trazan los rayos.

La ecuación (6.131), si contemplásemos al parámetro s como el tiempo, es la que describe el movimiento de una partícula de masa unitaria en un potencial de la forma

$$V = -\frac{1}{2}n^2 \tag{6.132}$$

y la condición (6.128) implica considerar ese movimiento cuando la energía total fuese

$$E = \frac{1}{2}\left|\frac{d\overline{r}}{dt}\right|^2 - \frac{1}{2}n^2 = 0. \tag{6.133}$$

La hamiltoniana de ese sistema sería

$$H = \frac{1}{2}\left(p_x^2 + p_y^2 + p_x^2\right) - \frac{1}{2}n^2 \tag{6.134}$$

y la ecuación de Hamilton Jacobi

$$\frac{1}{2}\left[\left(\frac{\partial G}{\partial x}\right)^2 + \left(\frac{\partial G}{\partial y}\right)^2 + \left(\frac{\partial G}{\partial z}\right)^2\right] - \frac{1}{2}n^2 = -\frac{\partial G}{\partial t}, \tag{6.135}$$

la que, al separar el tiempo, daría

$$\frac{1}{2}\left[\left(\frac{\partial W}{\partial x}\right)^2+\left(\frac{\partial W}{\partial y}\right)^2+\left(\frac{\partial W}{\partial z}\right)^2\right]-\frac{1}{2}n^2=\alpha_1. \tag{6.136}$$

Considerando que en esta última ecuación α_1 es la energía total y que, en el caso óptico, hemos restringido la energía total a ser cero, las soluciones de la ecuación de la eikonal (6.124), lo serán de ésta cuando $\alpha_1 = 0$. Vemos pues que los rayos de un sistema óptico en la aproximación de longitud de onda pequeña y las trayectorias de una partícula de masa unitaria en el potencial correspondiente y para energía total cero, son las mismas curvas, pues se determinan por las mismas ecuaciones y procedimientos matemáticos. Esta es la razón por la que, en el nivel de la óptica geométrica, una visión corpuscular de la luz y una visión ondulatoria son indistinguibles; los rayos de luz pueden verse como curvas perpendiculares a los frentes de onda cuando la longitud de onda es pequeña o como las trayectorias de unas partículas mecánicas de masa unitaria y energía total cero sujetas a un potencial de la forma $-(1/2)n^2$.

Ahora bien la ecuación de la eikonal es la aproximación a longitud de onda pequeña de una ecuación de ondas y es análoga a una ecuación de Hamilton Jacobi para un sistema mecánico; podemos entonces preguntarnos por la existencia de una ecuación de ondas tal que al tomar la aproximación de onda pequeña resulte, en general, la ecuación de Hamilton Jacobi para un sistema mecánico. Esto fue planteado por Hamilton mismo como una curiosidad, aunque no llegó a ninguna conclusión, en gran medida porque no veía ninguna razón física para construir una ecuación de ondas de esa naturaleza.

En cambio a mediados de los años veinte del siglo XX la situación era muy distinta. En el estudio de la interacción de la luz con la materia y del comportamiento de los átomos y las partículas subatómicas, se habían encontrado propiedades y fenómenos para los que las teorías clásicas, como la mecánica, generaban resultados que no concordaban con las observaciones experimentales. En adición a esto, se habían generado algunas explicaciones semiempíricas para las desviaciones que ocurrían en el mundo microscópico, aunque no una teoría general que los comprendiera; en particular los trabajos de Plank y Einsten en torno a fenómenos como la radiación de cuerpo negro y el efecto fotoeléctrico, habían asociado propiedades corpusculares a la luz en sus interacciones con la materia y, de Broglie, había dado explicación a la mecánica atómica o cuántica de Bohr y Somerfeld asociando un fenómeno ondulatorio a las partículas atómicas, ambas situaciones apuntaban hacia

una dualidad entre un comportamiento corpuscular y uno ondulatorio en la naturaleza de la luz y de las partículas atómicas o subatómicas. Ante esto renace la vieja idea de buscar un comportamiento ondulatorio para el cual la ecuación de Hamilton Jacobi resultara ser una aproximación de onda pequeña, al igual que la ecuación de la eikonal resulta serlo para la ecuación de ondas (6.105). Naturalmente en esta ocasión se tenía una motivación física real para hacerlo y una guía en las fórmulas semiempíricas de Plank, Einstein, Bohr, Somerfeld y de Broglie.

El propio de Broglie pensó en este asunto, pero fue Schroedinger quien, siguiendo un proceso heurístico, encontró una ecuación de ondas que cumplía con ese requisito y que daba explicación a los fenómenos atómicos, en especial a los espectros atómicos.

Para hacer ver esto seguiremos un camino inverso al que hubo de seguir Schroedinger, esto es partiremos de la, ya conocida, ecuación de ondas de Schroedinger y, por un proceso análogo al seguido para encontrar la ecuación de la eikonal a partir de la ecuación de ondas (6.105), le buscaremos una aproximación para longitudes de onda pequeña.

La ecuación de Schroedinger correspondiente a un sistema cuya hamiltoniana es $H(q,p)$ es

$$\hat{H}\psi = i\frac{h}{2\pi}\frac{\partial\psi}{\partial t}, \tag{6.137}$$

donde el operador hamiltoniano $\hat{H}$ se obtiene a partir de la hamiltoniana clásica substituyendo los impulsos p_i por los operadores

$$\hat{p_i} = -i\frac{h}{2\pi}\frac{\partial}{\partial q^i}. \tag{6.138}$$

Para una hamiltoniana de la forma general

$$H = \frac{1}{2}\sum_{ij} I^{ij}(q)p_i p_j + V(q), \tag{6.139}$$

donde I es el tensor de inercia en su expresión contravariante, tendremos que la ecuación de Schroedinger es

$$-\frac{h^2}{8\pi^2}\sum_{ij} I^{ij}(q)\frac{\partial^2\psi}{\partial q^j \partial q^i} + V(q)\psi = i\frac{h}{2\pi}\frac{\partial\psi}{\partial t}. \tag{6.140}$$

Busquemos una solución de la forma

$$\psi = A(q,t)e^{i\frac{2\pi}{h}S(q,t)} \tag{6.141}$$

con A y S funciones reales y que nos permite tener una fase y tomar a A y S como funciones reales. La longitud de onda[3] para esta solución será

$$\lambda = \frac{h}{|\bigtriangledown S|} \tag{6.142}$$

y la frecuencia angular

$$\omega = 2\pi\nu = \frac{2\pi}{\tau} = 2\pi\frac{\left|\frac{\partial S}{\partial t}\right|}{h}. \tag{6.143}$$

Al substituir (6.141) en la ecuación de Schroedinger obtenemos que

$$-\frac{1}{2}\sum_{ij} I^{ij}(q)\left\{\left[\frac{h^2}{4\pi^2}\frac{\partial^2 A}{\partial q^i \partial q^j} - A\frac{\partial S}{\partial q^i}\frac{\partial S}{\partial q^j}\right] + i\left[\frac{h}{\pi}\frac{\partial A}{\partial q^i}\frac{\partial S}{\partial q^j} + \frac{h}{2\pi}\frac{\partial^2 S}{\partial q^i \partial q^j}\right]\right\} \tag{6.144}$$

$$+V(q)A = -A\frac{\partial S}{\partial t} + i\frac{h}{2\pi}\frac{\partial A}{\partial t}, \tag{6.145}$$

esta expresión, al tomar las partes reales e imaginarias, proporciona dos ecuaciones a partir de las cuales se pueden obtener $A(q,t)$ y $S(q,t)$ y por tanto ψ.

Ahora en la aproximación de longitud de onda pequeña, la que (6.142) nos hace ver que se obtiene cuando $h \to 0$, (6.145) se reduce a

$$\frac{1}{2}\sum_{ij} I^{ij}(q)\frac{\partial S}{\partial q^i}\frac{\partial S}{\partial q^j} + V(q) = -\frac{\partial S}{\partial t}, \tag{6.146}$$

que es la ecuación de Hamilton Jacobi para el sistema clásico con hamiltoniana (6.139).

Es interesante ver que si se trata de una partícula en el espacio físico sujeta al potencial $V(x,y,z)$, entonces en el sentido clásico

$$|\bigtriangledown S| = \sqrt{\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial Z}\right)^2} = \sqrt{p_x^2 + p_y^2 + p_z^2} = p \tag{6.147}$$

y

$$\left|\frac{\partial S}{\partial t}\right| = |E|\,, \tag{6.148}$$

[3]Suponiendo una métrica euclidiana en el espacio de configuración y usando coordenadas cartesianas en él.

por lo que (6.142) y (6.143) quedan como

$$\lambda = \frac{h}{p} \tag{6.149}$$

y

$$|E| = \nu h, \tag{6.150}$$

que son las formulas de de Broglie y Plank respectivamente.

Vemos así que, de manera análoga a la relación existente entre la óptica geométrica y la óptica física, la mecánica clásica puede verse como una aproximación de onda pequeña de la mecánica cuántica o ondulatoria.

6.4. La función principal de Hamilton.

En el capítulo *** vimos que la mecánica clásica puede formularse a partir de un principio variacional, el de Hamilton. Se parte de la funcional de acción definida por

$$S = \int_\gamma L(q, \dot{q}, t)dt \tag{6.151}$$

donde γ es una curva $q(t)$ definida entre dos valores del tiempo t_0 y t_1. Dados dos puntos de la variedad de configuración q_0 y q_1, entre todas las curvas que pasan por q_0 al tiempo t_0 y por q_1 al tiempo t_1, aquella (o aquellas) en que la funcional de acción toma valores extremos, será (o serán) solución del sistema mecánico definido por la lagrangiana L. A partir de esto se define la función principal de Hamilton como el valor de la funcional de acción en esa curva extremal y la denotaremos [4] con la letra S.

La función principal de Hamilton dependerá de los dos puntos de la variedad de configuración q_0 y q_1 y de los tiempos t_0 y t_1, cuando la lagrangiana no depende explícitamente del tiempo la dependencia será únicamente de $t_1 - t_0$.

Para calcular la función principal de Hamilton se puede seguir el siguiente procedimiento.

1. Se resuelven las ecuaciones de Lagrange para obtener las soluciones $q^i = q^i(q_0, t_0, \dot{q}_0, t)$, donde q_0 y $\dot{q}_0$ son las condiciones iniciales al tiempo t_0.

[4] Nótese que se usa la misma notación para la funcional de acción y para la función principal de Hamilton, el contexto en que se use nos evitará confusiones.

En los casos en que se tenga más de una solución podrá haber más de un valor, pero no consideraremos esto aquí.

2. Se establecen las ecuaciones $q^i(q_0, t_0, \dot{q}_0, t_1) = q_1^i$ y de éstas se despejan los valores de las velocidades generalizadas iniciales $\dot{q}_0^i$, que al ser substituidas en las soluciones obtenidas en el paso anterior nos proporcionan $q^i = q^i(q_0, t_0, q_1, t_1, t)$, que es la ecuación de la trayectoria que pasa por q_0 a t_0 y por q_1 a t_1.

3. Por derivación se obtiene $\dot{q}^i = \dot{q}^i(q_0, t_0, q_1, t_1, t)$.

4. Finalmente se hace la integral

$$\int_{t_0}^{t_1} L(q(q_0, t_0, q_1, t_1, t), \dot{q}(q_0, t_0, q_1, t_1, t), t)dt, \tag{6.152}$$

que nos proporciona $S(q_0, t_0, q_1, t_1)$.

Este procedimiento se toma en ocasiones como una definición operacional de la función principal de Hamilton.

Vemos que el conocer las soluciones del sistema mecánico permite conocer la función principal de Hamilton. Lo inverso es también cierto, si se conoce la función principal de Hamilton se pueden encontrar las soluciones del sistema mecánico. Veamos esto último.

La función principal de Hamilton es

$$S(q_0, t_0, q_1, t_1) = \int_{t_0}^{t_1} L(q(q_0, t_0, q_1, t_1, t), \dot{q}(q_0, t_0, q_1, t_1, t), t)dt \tag{6.153}$$

donde q y $\dot{q}$ son soluciones de las ecuaciones de Lagrange. Veamos cual es el cambio en S cuando q_0 y q_1 cambian por δq_0 y δq_1 dejando t_0 y t_1 inalteradas.

$$\begin{aligned} \delta S &= \int_{t_0}^{t_1} \delta L(q, \dot{q}, t)dt \\ &= \int_{t_0}^{t_1} \sum_i \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i \right) dt, \end{aligned}$$

como $q(t)$ y $\dot{q}(t)$ son solución de las ecuaciones de Lagrange

$$\delta S = \int_{t_0}^{t_1} \sum_i \left(\dot{p}_i \delta q^i + p_i \delta \dot{q}^i \right) dt, \tag{6.154}$$

pero

$$\delta \dot{q}^i = \frac{d}{dt}(\delta q^i), \tag{6.155}$$

por lo que

$$\begin{aligned}\delta S &= \int_{t_0}^{t_1} \sum_i \frac{d}{dt}(p_i \delta q^i) dt = \sum_i \left[p_i \delta q^i\right]_{t_0}^{t_1} \\ &= \sum_i \left(p_{1i}\delta q_1^i - p_{0i}\delta q_0^i\right),\end{aligned}$$

donde hemos denotado por p_0 y p_1 a los valores de p a los tiempos t_0 y t_1 respectivamente.

Por otro lado tenemos que

$$\delta S = \sum_i \left(\frac{\partial S}{\partial q_0^i}\delta q_0^i + \frac{\partial S}{\partial q_1^i}\delta q_1^i\right), \tag{6.156}$$

por lo que, al ser δq_0^i y δq_1^i independientes, obtenemos

$$p_{1i} = \frac{\partial S}{\partial q_1^i} \quad y \quad p_{0i} = -\frac{\partial S}{\partial q_0^i}. \tag{6.157}$$

Contemplemos ahora S como una función de un punto inicial q_0 a tiempo t_0 y uno final variable q a tiempo t, esto no es más que llamar q a q_1 y t a t_1 una vez construida la función. Tomemos la segunda de las ecuaciones (6.157). p_{0i} es una constante y el lado derecho es una función de q y t por tanto si conocemos S podemos despegar y obtener las q^i como funciones de q_0, t_0, p_0 y t, estas serán las trayectorias del sistema mecánico para condiciones iniciales q_0, p_0 a t_0. Por otro lado si tomamos las primeras de las ecuaciones (6.157) y en ellas substituimos la recién obtenida $q(t)$ obtenemos p_i como función de q_0, t_0, p_0 y t, lo que completa la solución del sistema mecánico. Así, si se conoce la solución de un sistema mecánico, el procedimiento indicado arriba nos permite calcular la función principal de Hamilton y si se conoce la función principal de Hamilton las ecuaciones (6.157) nos proporcionan la solución de las ecuaciones de movimiento en el espacio fase. Nótese que el procedimiento sustentado en las ecuaciones (6.157) para encontrar la solución de las ecuaciones de movimiento es el mismo que se utiliza para encontrarlas a partir de una solución completa de la ecuación de Hamilton Jacobi del sistema si identificamos a las q_0^i con las α^i y a las p_{0i} con las β_i de la ecuación 6.8. Todo esto nos induce a pensar que la función principal de Hamilton es una solución completa de la ecuación de Hamilton Jacobi, exploremos esta posibilidad con detenimiento.

Calculemos la derivada de S con respecto a t_1; por definición

$$S(q_0, t_0, q_1, t_1) = \int_{t_0}^{t_1} L\left(q(q_0, t_0, q_1, t_1, t), \dot{q}(q_0, t_0, q_1, t_1, t), t\right) dt \tag{6.158}$$

donde $q(q_0,t_0,q_1,t_1,t)$ es la trayectoria que a t_0 pasa por q_0 y a t_1 pasa por q_1, $\dot{q}(q_0,t_0,q_1,t_1,t)$ denota a la velocidad generalizada sobre esa trayectoria. Al hacer un incremento δt_1 en t_1 se obtiene un incremento en S dado por

$$\begin{aligned}
\delta S &= \int_{t_0}^{t_1+\delta t_1} L\left(q(q_0,t_0,q_1,t_1+\delta t_1,t),\dot{q}(q_0,t_0,q_1,t_1+\delta t_1,t),t\right)dt - \\
&- \int_{t_0}^{t_1} L\left(q(q_0,t_0,q_1,t_1,t),\dot{q}(q_0,t_0,q_1,t_1,t),t\right)dt = \\
&= \int_{t_0}^{t_1} L\left(q(q_0,t_0,q_1,t_1+\delta t_1,t),\dot{q}(q_0,t_0,q_1,t_1+\delta t_1,t),t\right)dt - \\
&- \int_{t_0}^{t_1} L\left(q(q_0,t_0,q_1,t_1,t),\dot{q}(q_0,t_0,q_1,t_1,t),t\right)dt + \\
&+ \int_{t_1}^{t_1+\delta t_1} L\left(q(q_0,t_0,q_1,t_1+\delta t_1,t),\dot{q}(q_0,t_0,q_1,t_1+\delta t_1,t),t\right)dt.
\end{aligned}$$

Al calcular esta expresión a primer orden en δt_1 se obtiene

$$\begin{aligned}
\delta S &= \delta t_1 \int_{t_0}^{t_1} \frac{\partial L\left(q(q_0,t_0,q_1,t_1,t),\dot{q}(q_0,t_0,q_1,t_1,t),t\right)}{\partial t_1}dt + \\
&+ \int_{t_1}^{t_1+\delta t_1} L\left(q(q_0,t_0,q_1,t_1,t),\dot{q}(q_0,t_0,q_1,t_1,t),t\right)dt = \\
&= \delta t_1 \int_{t_0}^{t_1} \left[\sum_i \left(\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial t_1} + \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial t_1}\right)\right]dt + \\
&+ \int_{t_1}^{t_1+\delta t_1} L\left(q(q_0,t_0,q_1,t_1,t),\dot{q}(q_0,t_0,q_1,t_1,t),t\right)dt = \\
&= \delta t_1 \left[\int_{t_0}^{t_1} \sum_i \left(\frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial t_1} + \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial t_1}\right)dt + L(q,\dot{q},t)\right].
\end{aligned}$$

Ahora todo esto se hace sobre una trayectoria del sistema mecánico, por lo que se deben cumplir las ecuaciones de Lagrange, entonces

$$\frac{\partial L}{\partial q^i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \tag{6.159}$$

y además tenemos que

$$\frac{\partial \dot{q}_i}{\partial t_1} = \frac{d}{dt}\left(\frac{\partial q^i}{\partial t_1}\right), \tag{6.160}$$

por lo que

$$\begin{aligned}
\delta S &= \delta t_1 \left[\int_{t_0}^{t_1} \sum_i \left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \frac{\partial q^i}{\partial t_1} + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \left(\frac{\partial q^i}{\partial t_1} \right) \right) dt + L(q_1, \dot{q}_1, t_1) \right] = \\
&= \delta t_1 \left[\int_{t_0}^{t_1} \frac{d}{dt} \sum_i \left(\frac{\partial L}{\partial \dot{q}^i} \frac{\partial q^i}{\partial t_1} \right) dt + L(q_1, \dot{q}_1, t_1) \right] = \\
&= \delta t_1 \left[\sum_i \left(\frac{\partial L}{\partial \dot{q}^i} \frac{\partial q^i}{\partial t_1} \right) \Bigg|_{t_0}^{t_1} + L(q_1, \dot{q}_1, t_1) \right] = \\
&= \delta t_1 \left[\sum_i \left(\frac{\partial L}{\partial \dot{q}^i} \Big|_{t_1} \frac{\partial q^i}{\partial t_1} \Big|_{t_1} - \frac{\partial L}{\partial \dot{q}^i} \Big|_{t_0} \frac{\partial q^i}{\partial t_1} \Big|_{t_0} \right) + L(q_1, \dot{q}_1, t_1) \right] = \\
&= \delta t_1 \left[\sum_i \left(p_{1i} \frac{\partial q^i}{\partial t_1} \Big|_{t_1} - p_{0i} \frac{\partial q^i}{\partial t_1} \Big|_{t_0} \right) + L(q_1, \dot{q}_1, t_1) \right]. \qquad (6.161)
\end{aligned}$$

Ahora a primer orden

$$\frac{\partial q^i}{\partial t_1} \Big|_{t_0} \delta t_1 = q^i(q_0, t_0, q_1, t_1 + \delta t_1, t_0) - q^i(q_0, t_0, q_1, t_1, t_0), \qquad (6.162)$$

pero el primer término de la derecha es q_0^i y el segundo también por lo que

$$\frac{\partial q^i}{\partial t_1} \Big|_{t_0} = 0. \qquad (6.163)$$

Por otro lado, también a primer orden

$$\frac{\partial q^i}{\partial t_1} \Big|_{t_1} \delta t_1 = q^i(q_0, t_0, q_1, t_1 + \delta t_1, t_1) - q^i(q_0, t_0, q_1, t_1, t_1) \qquad (6.164)$$

y

$$\dot{q}_1^i \delta t_1 = q^i(q_0, t_0, q_1, t_1, t_1 + \delta t_1) - q^i(q_0, t_0, q_1, t_1, t_1) \qquad (6.165)$$

pero el primer término de la derecha de la primera ecuación es el valor de la coordenada q^i al tiempo t_1 sobre la trayectoria que partiendo de q_0 al tiempo t_0 pasa por q_1 al tiempo $t_1 + \delta t_1$, mientras que en la segunda ecuación el primer término de la derecha es el valor de la coordenada q^i al tiempo $t_1 + \delta t_1$ sobre la trayectoria que pasa por q_0 al tiempo t_0 y por q_1 al tiempo t_1, por lo que a primer orden

$$q^i(q_0, t_0, q_1, t_1 + \delta t_1, t_1) = -q^i(q_0, t_0, q_1, t_1, t_1 + \delta t_1) \qquad (6.166)$$

y encontramos que

$$\left.\frac{\partial q^i}{\partial t_1}\right|_{t_1} = -\dot{q}_1^i. \tag{6.167}$$

Al substituir todo esto en (6.161) encontramos que

$$\frac{\partial S}{\partial t_1} = -\sum_i p_{1i}\dot{q}_1^i + L(q_1, \dot{q}_1, t_1), \tag{6.168}$$

pero esta última expresión es el negativo de la función hamiltoniana valuada en q_1, p_1 y t_1 esto es

$$\frac{\partial S}{\partial t_1} = -H(q_1, p_1, t_1). \tag{6.169}$$

Esta expresión, al tomar en consideración la primera de las ecuaciones (6.157), nos dice que

$$\frac{\partial S}{\partial t_1} = -H\left(q_1, \frac{\partial S}{\partial q_1}, t_1\right); \tag{6.170}$$

una vez hechos los cálculos podemos cambiar de notación y llamar, como hicimos antes, a q_1 q y a t_1 t y encontramos que la función principal de Hamilton cumple con la ecuación de Hamilton Jacobi. En este caso las n constantes q_0 representan a las que llamamos α y las p_0 a las que llamamos β cuando tratamos con soluciones de la ecuación de Hamilton Jacobi en general.

Anteriormente encontramos una solución a la ecuación de Hamilton Jacobi para una partícula libre en dos dimensiones (6.34), calculemos ahora la función principal de Hamilton para este sistema.

La función la lagrangiana es

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) \tag{6.171}$$

y la solución en términos de la posición y la velocidad inicial

$$\begin{aligned} x &= v_{0x}t + x_0, \\ y &= v_{0y}t + y_0, \end{aligned}$$

donde hemos supuesto que $t_0 = 0$. Esta solución pasa por (x_0, y_0) a $t = 0$, si deseamos que pase por (x_1, y_1) a $t = t_1$ debemos tener que

$$\begin{aligned} x_1 &= v_{0x}t_1 + x_0, \\ y_1 &= v_{0y}t_1 + y_0, \end{aligned}$$

y por tanto

$$\begin{aligned} v_{0x} &= \frac{x_1 - x_0}{t_1}, \\ v_{0y} &= \frac{y_1 - y_0}{t_1}. \end{aligned}$$

La lagrangiana evaluada sobre la solución será

$$L(x_0, y_0, x_1, y_1, t_1, t) = \frac{1}{2} m \frac{(x_1 - x_0)^2 + (y_1 - y_0)^2}{t_1^2}, \tag{6.172}$$

por tanto

$$S(x_0, y_0, x_1, y_1, t_1) = \int_0^{t_1} L dt = \frac{1}{2} m \frac{(x_1 - x_0)^2 + (y_1 - y_0)^2}{t_1} \tag{6.173}$$

y eliminando, por razones de notación, el subíndice 1 tendremos

$$S(x_0, y_0, x, y, t) = \frac{1}{2} m \frac{(x - x_0)^2 + (y - y_0)^2}{t}. \tag{6.174}$$

De aquí y usando (6.157)

$$\begin{aligned} p_{0x} &= m \frac{x - x_0}{t}, \\ p_{0y} &= m \frac{y - y_0}{t} \end{aligned}$$

y despejando

$$\begin{aligned} x &= \frac{p_{0x}}{m} t + x_0, \\ y &= \frac{p_{0y}}{m} t + y_0. \end{aligned}$$

Es fácil hacer ver que (6.174) satisface la ecuación de Hamilton Jacobi para este caso. Vemos que es una solución distinta a la que obtenemos por separación de variables (6.34), es notoria la diferencia en la forma de la dependencia en el tiempo.

Las curvas de S constante a un tiempo fijo son, en este caso, círculos concéntricos con centro en (x_0, y_0), como lo son las curvas de fase constante de ondas circulares, y las trayectorias del sistema determinadas por esta solución conforman un haz de rectas que pasan por ese punto. Esto contrasta con la solución de la ecuación de Hamilton Jacobi para la partícula libre que

encontramos por separación de variables (6.67), donde las curvas de nivel son rectas paralelas, que semejan frentes de ondas planas, y determinan trayectorias que conforman un haz de rectas paralelas perpendiculares a las de nivel.

Tomemos otro ejemplo, el tiro parabólico. Utilizando un sistema cartesiano de coordenadas (x, y) tendremos una lagrangiana

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy. \tag{6.175}$$

La solución en términos de condiciones iniciales a $t = 0$ es

$$\begin{aligned} x &= v_{0x}t + x_0, \\ y &= -\frac{1}{2}gt^2 + v_{0y}t + y_0. \end{aligned} \tag{6.176}$$

En vez usar las condiciones iniciales (v_{0x}, v_{0y}) pongamos la solución en términos de t_1, x_1 y y_1, esto es de un punto y el tiempo en que la partícula debe pasar por ese punto. Entonces

$$\begin{aligned} x_1 &= v_{0x}t_1 + x_0, \\ y_1 &= -\frac{1}{2}gt_1^2 + v_{0y}t_1 + y_0, \end{aligned}$$

despejando las velocidades y substituyendo en (6.176) obtenemos

$$\begin{aligned} x &= \frac{x_1 - x_0}{t_1}t + x_0, \\ y &= -\frac{1}{2}gt^2 + \frac{y_1 - y_0 + \frac{1}{2}gt_1^2}{t_1}t + y_0 \end{aligned}$$

y

$$\begin{aligned} \dot{x} &= \frac{x_1 - x_0}{t_1}, \\ \dot{y} &= -gt + \frac{y_1 - y_0 + \frac{1}{2}gt_1^2}{t_1}, \end{aligned}$$

La lagrangiana evaluada sobre esta solución es

$$L = \frac{1}{2}m\left[\frac{(x_1 - x_0)^2}{t_1^2} + \frac{\left(y_1 + \frac{1}{2}gt_1^2 - y_0\right)^2}{t_1^2} - 2gy_0 - \frac{4g}{t_1}\left(y_1 + \frac{1}{2}gt_1^2 - y_0\right)t + 2gt^2\right], \tag{6.177}$$

por definición

$$S = \int_0^{t_1} L(x_0, y_0, x_1, y_1, t_1, t)dt, \tag{6.178}$$

al hacer esta integral y llamar, una vez hecha la integral, t a t_1, x a x_1 y y a y_1 obtenemos

$$S(x_0, y_0, x, y, t) = \frac{m}{2t}\left[(x - x_0)^2 + (y - y_0)^2\right] - \frac{mg}{2}(y + y_0)t - \frac{m}{24}g^2t^3. \tag{6.179}$$

A partir de (6.157) obtenemos

$$\begin{aligned} p_{0x} &= \frac{m}{t}(x - x_0), \\ p_{0y} &= \frac{m}{t}(y - y_0) + \frac{mg}{2}t, \\ p_x &= \frac{m}{t}(x - x_0), \\ p_y &= \frac{m}{t}(y - y_0) - \frac{mg}{2}t \end{aligned}$$

se puede ver que al despejar x y y de las dos primeras ecuaciones obtenemos la solución del sistema y que las dos últimas nos proporcionan la evolución temporal de los impulsos. Por otro lado es fácil comprobar que (6.179) es solución de la ecuación de Hamilton Jacobi del sistema

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right] + mgy = -\frac{\partial S}{\partial t}. \tag{6.180}$$

Recordamos que al resolver esta ecuación por el método de separación de variables obtenemos (6.78)

$$S = \alpha_2 x \mp \frac{1}{3m^2g}\left(2m\alpha_1 - \alpha_2^2 - 2m^2gy\right)^{\frac{3}{2}} - \alpha_1 t, \tag{6.181}$$

donde α_1 y α_2 son contantes arbitrarias. Podemos ver que se trata de soluciones distintas, aunque las dos proporcionan soluciones para el mismo sistema mecánico, de hecho puede haber muchas otras soluciones distintas y todas ellas proporcionarían soluciones del tiro parabólico. Históricamente en primer lugar Hamilton encontró las relaciones (6.157) entre la función principal y las soluciones de las ecuaciones de Hamilton y posteriormente Jacobi encontró que cualquier solución completa de la ecuación de Hamilton Jacobi generaba las soluciones de los sistemas mecánicos. En muchas ocasiones se define la función principal de Hamilton como aquella función que es solución

de la ecuación de Hamilton Jacobi y en el caso en que el tiempo sea separable a la parte de la solución que depende únicamente de las coordenadas se le llama función característica de Hamilton, en este texto hemos preferido limitar la nomenclatura a la función definida por (6.153).

6.5. La ecuación de Hamilton Jacobi y transformaciones canónicas.

En el capítulo anterior encontramos un importante teorema que nos dice que el flujo de un sistema hamiltoniano es, para cada valor del tiempo, un mapeo del espacio fase en si mismo que deja invariante a la forma simpléctica. Esto nos hace ver que si interpretamos el flujo como una transformación de coordenadas dependiente del tiempo, esta será una transformación canónica. Así si tenemos un sistema canónico de coordenadas (q_0, p_0) y en él el flujo del sistema se expresa por

$$q^i = q^i(q_0, p_0, t), \tag{6.182}$$

$$p_i = p_i(q_0, p_0, t), \tag{6.183}$$

podemos interpretar estas fórmulas como una transformación canónica de coordenadas dependiente del tiempo que lleva de un sistema de coordenadas (q_0, p_0) a otro sistema canónico (q, p). Este último sistema será uno que se mueve acompañando al flujo, por lo que en él las soluciones de las ecuaciones de movimiento deben ser

$$\begin{aligned} q^i &= \text{cst.} \\ p_i &= \text{cst.} \end{aligned}$$

y por tanto la hamiltoniana, en ese sistema canónico en movimiento, debe ser la función constante, que sin perdida de la generalidad, podemos considerar como cero. Esto es la hamiltoniana $H = H(q_0, p_0, t)$ en el sistema (q_0, p_0) se transforma en $H^{'} = 0$ en el sistema (q, p).

Vemos que el conocimiento de esa transformación canónica de coordenadas que lleva a una nueva hamiltoniana cero, nos proporciona las soluciones de las ecuaciones de movimiento, por lo que tener un procedimiento que nos permita encontrarla sería de gran utilidad. Ahora bien, esa transformación es canónica por lo que debe existir una función que la genere, así que únicamente requerimos encontrar esa función generadora para obtener la solución general del sistema mecánico. Llamemos a dicha función generadora F y

supongamos que es de tipo uno, esto es $F = F(q_0, q, t)$, donde recordamos que q_0 son las coordenadas originales, q las nuevas coordenadas y que las transformación es dependiente del tiempo por lo que tendremos que

$$p_{0i} = -\frac{\partial F}{\partial q_0^i}, \quad p_i = \frac{\partial F}{\partial q^i} \quad y \quad H^{'} = H + \frac{\partial F}{\partial t}, \tag{6.184}$$

a lo que hay que adicionar la condición

$$\left| \frac{\partial^2 F}{\partial q_0^i \partial q} \right| \neq 0. \tag{6.185}$$

Pero, para la transformación que buscamos, $H^{'}$ debe ser cero, o sea que

$$H(q_0, p_0, t) + \frac{\partial F}{\partial t} = 0, \tag{6.186}$$

que al substituir p_{0i} por $\partial F / \partial q_0^i$ nos dice que

$$H(q_0^1, \ldots, q_0^n, \frac{\partial F}{\partial q_0^1}, \ldots, \frac{\partial F}{\partial q_0^n}, t) = -\frac{\partial F}{\partial t}, \tag{6.187}$$

que no es otra que la ecuación de Hamilton Jacobi, cuyas soluciones completas, o sea que contienen n constantes arbitrarias serán de la forma $F(q_0, q, t)$ ya que las q^i son n constantes. Esto es, la ecuación de Hamilton Jacobi puede interpretarse como una ecuación cuyas soluciones completas determinan funciones generadoras de transformaciones canónicas que llevan, del sistema original de coordenadas canónicas, a uno nuevo en el que la hamiltoniana es cero y, por tanto, las soluciones de las ecuaciones de movimiento son todas constantes. Una cualquiere de esas transformaciones canónicas de coordenadas, vista como mapeo del espacio fase en si mismo, es el flujo del sistema mecánico.

Desde este punto de vista la ecuación de Hamilton Jacobi proporciona una función generadora de transformaciones canónica, las α introducidas en (6.4) son las nuevas coordenadas y las β de (6.8) son los nuevos impulsos, ambas cantidades deben ser constantes porque así se requirió. El hecho de que existan distintas soluciones de la ecuación de Hamilton Jacobi nos indica que existen muchos sistemas de coordenadas canónicas que siguen al flujo, de hecho por cada sistema de coordenadas canónicas hay uno, o dicho de otra forma, el flujo se expresa de distinta manera en distintos sistemas de coordenadas.

Capítulo 7

Sistemas mecánicos integrables.

En los capítulos anteriores hemos desarrollado diversas formulaciones de la mecánica clásica. Esto nos ha permitido contemplar las leyes naturales que gobiernan el movimiento desde diversas perspectivas físicas y matemáticas, y nos ha proporcionado una mayor capacidad de entenderlas, de extraer propiedades generales del movimiento y de generar métodos para resolver casos concretos. Desde la visión newtoniana hasta la ecuación de Hamilton Jacobi hemos contemplado el basto panorama de los principios y causas del movimiento, pero poco hemos dicho sobre el movimiento mismo, esto es, sobre las formas o tipos de movimiento posibles, sobre su posible clasificación, sobre sus propiedades. Hemos si estudiado, a guisa de ejemplo o por su propia importancia, algunos movimientos concretos como el sistema kepleriano o el péndulo esférico, pero no hemos dicho nada sobre las características generales del movimiento.

En este capítulo iniciaremos este estudio, naturalmente lo haremos en el espacio fase, pues al ser este un espacio de estados de los sistemas mecánicos y poder expresar en él las ecuaciones de movimiento como un campo vectorial hamiltoniano, las propiedades del campo se reflejan de manera natural en propiedades de su flujo, que no son otras que las del movimiento físico.

7.1. Sistemas integrables.

A lo largo de nuestro estudio de la mecánica hemos destacado la importancia que tienen las constantes de movimiento y, en especial, las integrales de movimiento. La presencia de una de estas cantidades conservadas nos permite eliminar una de las $2n$ ecuaciones de movimiento, lo que simplifica el estudio del sistema. El estado físico de un sistema mecánico de n grados de

libertad está determinado por $2n$ cantidades, por lo que, para solucionarlo totalmente requerimos conocer $2n$ constantes de movimiento independientes que no podrán ser todas ellas integrales de movimiento, pues éstas determinarían totalmente el estado físico del sistema independientemente del tiempo y no existiría el movimiento, esto es, algunas de las constantes de movimiento son necesariamente funciones explicitas del tiempo. Así las dificultadoes de resolver totalmente un sistema mecánico son totalmente equivalentes a las de encontrar un conjunto de $2n$ constantes de movimiento independientes.

Ahora bien, resulta que es posible llegar a conocer totalmente un sistema mecánico cuando se conocen únicamente n integrales de movimiento independientes si, en adición, el paréntesis de Poisson entre cualesquiera dos de ellas es cero; propiedad, esta última, a la que se le llama involución. Destacamos que sistemas como el de una partícula en un campo central, el péndulo esférico y otros que hemos reducido a cuadraturas, pertenecen a esta clase de sistemas y en ellos nos bastó tener n integrales de movimiento independientes y en involución para llegar conocer las otras n constantes de movimiento, éstas surgieron en la forma de las constantes de integración durante el proceso de reducción a cuadraturas. Pasemos pues a estudiar la clase de sistemas mecánicos que poseen tantas integrales de movimiento linealmente independientes y en involución como grados de libertad, sistemas mecánicos a los que llamaremos integrables.

Cuando tenemos un sistema integrable de n grados de libertad, la independencia lineal de sus integrales de movimiento y el que éstas estén en involución, nos permitiría considerarlas como una posible parte de un sistema canónico de coordenadas, ya sean las coordenadas o los impulsos, si es que fuera posible complementarlas con otras que fueran sus canónicas conjugadas. Si esto se logra tendríamos un sistema canónico en el que la mitad de las coordenadas, digamos los impulsos, permanecerían constantes a lo largo del movimiento, por lo que, la hamiltoniana del sistema dependería únicamente de esas coordenadas, esto es, todas las otras deberían ser ignorables y las ecuaciones de Hamilton serían muy simples; unas, las derivadas temporales de las integrales de movimiento, serían cero y las derivadas con respecto al tiempo de las otras coordenadas serían constantes, por lo que en esas coordenadas quedaría integrado el sistema. Estudiemos entonces la posibilidad de que exista un sistema canónico de coordenadas con estas características.

Consideremos un sistema hamiltoniano integrable con hamiltoniana $H(q,p)$[1] en algunas coordenadas canónicas (q,p), esto es, existen n integrales de mo-

[1] En este capítulo consideraremos únicamente hamiltonianas independientes del tiempo.

vimiento, $I_i(q,p)$, linealmente independientes y en involución, o sea

$$\{I_i, I_j\} = 0 \quad \text{para toda } i \text{ y } j \tag{7.1}$$

Como decíamos, la involución nos permite pensar en las I como candidatas a ser los impulsos[2] de algún sistema canónico de coordenadas, pero no sabemos si en general éste realmente exista. Pasemos a ver esta cuestión.

Tenemos n integrales de movimiento y $2n$ coordenadas (q,p), podemos, dada la independencia lineal, despejar n de las coordenadas en función de las otras n y de las integrales de movimiento, hagamos esto para los impulsos p

$$p_i = \pi_i(q, I). \tag{7.2}$$

Tomemos ahora la identidad

$$I_i \equiv I_i(q, \pi(q, I)), \tag{7.3}$$

la que al derivar con respecto a q^k y considerar a las cantidades I_i del término izquierdo como independientes nos dará

$$0 = \frac{\partial I_i}{\partial q^k} + \sum_l \frac{\partial I_i}{\partial \pi_l}\frac{\partial \pi_l}{\partial q^k}, \tag{7.4}$$

multiplicando por $\partial I_j/\partial p_k$ y sumando sobre k obtenemos

$$0 = \sum_k \frac{\partial I_i}{\partial q^k}\frac{\partial I_j}{\partial p_k} + \sum_{kl} \frac{\partial I_i}{\partial \pi_l}\frac{\partial I_j}{\partial p_k}\frac{\partial \pi_l}{\partial q^k}. \tag{7.5}$$

Podemos ahora intercambiar el orden en que hacemos estas operaciones para obtener

$$0 = \sum_k \frac{\partial I_i}{\partial p_k}\frac{\partial I_j}{\partial q^k} + \sum_{kl} \frac{\partial I_j}{\partial \pi_l}\frac{\partial I_i}{\partial p_k}\frac{\partial \pi_l}{\partial q^k}. \tag{7.6}$$

En el segundo término de la derecha de esta ecuación se suma sobre los índices l y k por lo que son intercambiables y podemos escribir la ecuación como

$$0 = \sum_k \frac{\partial I_i}{\partial p_k}\frac{\partial I_j}{\partial q^k} + \sum_{kl} \frac{\partial I_j}{\partial \pi_k}\frac{\partial I_i}{\partial p_l}\frac{\partial \pi_k}{\partial q^l}, \tag{7.7}$$

[2]Podrían ser los impulsos o las coordenadas, es lo mismo. En un sistema arbitrario de coordenadas canónicas simplemente la mitad es conjugada de la otra mitad y no son necesariamente coordenadas para la variedad de configuración unas e impulsos las otras. Aquí seguiremos el uso de llamar a las integrales de movimiento impulsos.

pero, por tratarse de notación, las parciales con respecto a π y a p son lo mismo, por lo que podemos escribir esta ecuación como

$$0 = \sum_k \frac{\partial I_i}{\partial p_k}\frac{\partial I_j}{\partial q^k} + \sum_{kl} \frac{\partial I_j}{\partial p_k}\frac{\partial I_i}{\partial \pi_l}\frac{\partial \pi_k}{\partial q^l}, \tag{7.8}$$

al restar de (7.5) (7.8) obtenemos que

$$\{I_i, I_j\} = \sum_{kl} \frac{\partial I_i}{\partial \pi_l}\frac{\partial I_j}{\partial p_k}\left[\frac{\partial \pi_k}{\partial q^l} - \frac{\partial \pi_l}{\partial q^k}\right] \tag{7.9}$$

esta última ecuación puede verse como un sistema de ecuaciones que nos permite obtener las cantidades

$$\frac{\partial \pi_k}{\partial q^l} - \frac{\partial \pi_l}{\partial q^k} \tag{7.10}$$

conociendo los paréntesis de Poisson de las I. Si estos paréntesis son cero, esto es, si las I están en involución, entonces esas cantidades deben ser cero, por ser los coeficientes del sistema de ecuaciones independientes, y por tanto

$$\frac{\partial \pi_k}{\partial q^l} = \frac{\partial \pi_l}{\partial q^k}. \tag{7.11}$$

Esta expresión nos hace ver que si consideramos a las I como parámetros independientes la expresión

$$\sum_i \pi_i(q, I) dq^i \tag{7.12}$$

es integrable, por lo que existe una función, calculable por cuadraturas, $F(q, I)$ tal que

$$p_i = \pi_i(q, I) = \frac{\partial F(q, I)}{\partial q^i}. \tag{7.13}$$

Ahora $F(q, I)$ puede ser utilizada como función generadora de transformaciones canónicas de tipo dos y al complementar (7.13) con

$$Q^i = \frac{\partial F(q, I)}{\partial I_i}, \tag{7.14}$$

obtendremos una transformación canónica de coordenadas

$$(q, p) \longleftrightarrow (Q, I), \tag{7.15}$$

que nos lleva a un sistema canónico de coordenadas donde los impulsos son las integrales de movimiento en involución; en estas coordenadas la hamiltoniana sería únicamente función de las I

$$H = H(I_1, \ldots, I_n) \tag{7.16}$$

y las ecuaciones de movimiento

$$\dot{Q}^i = \frac{\partial H}{\partial I_i} = \text{cst.}, \tag{7.17}$$

$$\dot{I}_i = 0. \tag{7.18}$$

Estas ecuaciones son fácilmente integrables y aplicando la transformación a la inversa podemos tener la solución del sistema expresada en las coordenadas originales.

Así hemos demostrado que la independencia lineal y la involución implican la existencia del sistema canónico de coordenadas que buscábamos y encontramos un procedimiento explícito que reduce a cuadraturas la solución del sistema, o sea, todo sistema integrable es reducible a cuadraturas. Este resultado es conocido como teorema de Liuville.

Desde una perspectiva global este resultado podemos verlo de la siguiente manera.

Las n integrales de movimiento son funciones del espacio fase en los reales, tomemos el campo vectorial hamiltoniano correspondiente a cada una de ellas V_{I_i}, recordando la ecuación *** tendremos que

$$\gamma\left(V_{I_i}, V_{I_j}\right) = \{I_i, I_j\} = 0. \tag{7.19}$$

Como las I son diferenciables y linealmente independientes, para cualquier eneada de constantes $(c_1, \ldots, c_n)$ la hipersuperficie de nivel $I_i = c_i \quad i = 1, \ldots, n$ es una subvariedad diferencial de dimensión n del espacio fase que denotaremos por M_c. Sobre esta subvariedad, esto es en un punto de M_c, los n vectores de los n campos vectoriales V_{I_i} son tangentes a M_c, pues al estar las I en involución los flujos hamiltonianos de cada una de ellas deja invariantes a las demás y por tanto a M_c. La independencia lineal de las I nos hace ver que estos n vectores son independientes y, por tanto, forman una base del espacio tangente a M_c en el punto en cuestión. Ahora, de (7.19) vemos que, si tenemos dos campos vectoriales cualesquiera, X y Y, tangentes a M_C, tendremos que sobre esta subvaridedad

$$\gamma(X, Y) = 0, \tag{7.20}$$

ya que, en todo punto de M_c, X y Y serán una combinación lineal de la base formada por los campos V_{I_i}. En general, a una subvariedad de una variedad simpléctica tal que la función que resulta al aplicar la forma simpléctica a dos campos tangentes a ella vale cero sobre dicha subvariedad, se le llama subvariedad nula o lagrangiana, en nuestro caso M_c es una subvariedad lagrangiana del espacio fase.

Ahora γ es la diferencial exterior de la uno forma diferencial ω, por lo que al restringirnos a M_c tendremos que la diferencial de ω es cero y entonces ω, sobre M_c, es exacta y debe existir una función $F_c : M_c \to R$ tal que $\omega = dF_c$, esta función será distinta para cada conjunto de valores $\{c_i\}$, como también será distinta la subvariedad, por lo que podemos verla como una colección de funciones parametrizada por estos valores. Podemos ahora utilizar a $F_c(q)$ como una función generadora de transformaciones canónicas $F(q, c)$ o, lo que es lo mismo, $F(q, I)$

Si tenemos un sistema local de coordenadas canónicas (q, p) las integrales se expresarán como $I_i(q, p)$ y las subvariedades lagrangianas M_c estará determinada por $I_i(q, p) = c_i$, por lo que, podremos utilizar a $F_c(q)$ como una función generadora de transformaciones canónicas, esta transformación nos lleva a unas coordenadas canónicas en las que las I son los impulsos. Naturalmente podremos escribir que

$$p_i = \pi_i(q, c), \tag{7.21}$$

y entonces

$$dF_c(q) = \omega = \sum_i p_i dq^i = \sum_i \pi_i(q, c) dq^i \tag{7.22}$$

que no es otra cosa que (7.13), con esto recuperamos el proceso hecho en coordenadas, claro está que ahora tenemos una demostración global de la existencia de F_c, esto es, la función está bien definida en toda la subvariedad lagrangiana M_c para cada conjunto de valores de c.

Cuando tenemos un sistema integrable siempre podemos tomar a la función hamiltoniana como una de las integrales de movimiento en involución, por lo que el campo hamiltoniano del sistema será tangente a las subvariedades lagrangianas M_c. Esto nos muestra que el flujo del sistema estructura al espacio fase en estas subvariedades M_c, pues toda trayectoría o linea de flujo del sistema permanecerá sobre alguna de ellas, la que le corresponda según el valor que tengan las integrales de movimiento para las condiciones iniciales correspondientes a esa trayectoria. Como hay un sinnúmero de trayectorias que comparten el mismo conjunto de valores de las integrales de movimiento, aunque corresponden a valores distintos de las otras n condiciones iniciales,

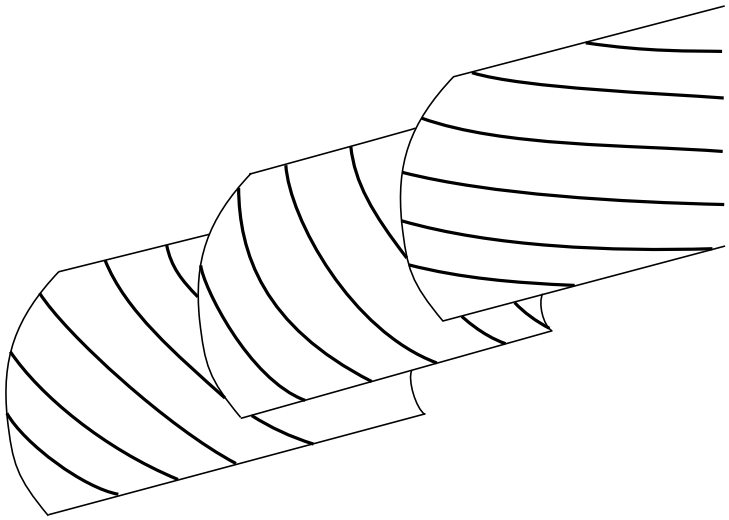


Figura 7.1: Las trayectorias forman haces que delinean subvariedades M_c de dimensión n.

las subvariedades M_c parecen formadas por las trayectorias, como si el espacio fase estuviera estructurado en capas de dimensión n (las subvariedades M_c) y, a su vez, cada una de estas capas quedará constituida por un haz de trayectorias del sistema; en la figura (7.1) se ilustra esto para el caso de dos grados de libertad, donde el espacio fase tiene cuatro dimensiones y las subvariedades M_c dos.

El espacio fase queda así estructurado en foliaciones por subvariedades de dimensión n que a su vez están compuestas de trayectorias. Naturalmente como las funciones que definen a las integrales de movimiento son bien comportadas, tendremos que las variedades M_c estarán bien definidas excepto para un conjunto de medida cero de eneadas c_i; para los valores en que ocurra una de estas singularidades, M_c será una separatriz, esto es, una hipersuperficie que separa una foliación local del espacio fase de otra. Tomemos como caso ilustrativo de esto al péndulo simple.

El péndulo simple tiene un grado de libertad, por lo que es integrable. Su espacio fase es un cilindro y las trayectorias generan sobre dicho cilindro la configuración que se muestra en la figura 7.2. Por tratarse de un sistema de un grado de libertad las subvariedades M_c serán de dimensión uno, o sea, son curvas cerradas que constituyen foliaciones locales y coinciden con las trayectorias; se presentan como singularidades los dos puntos de equilibrio y las separatrices que dividen los movimientos de libración de los de rotación. Si

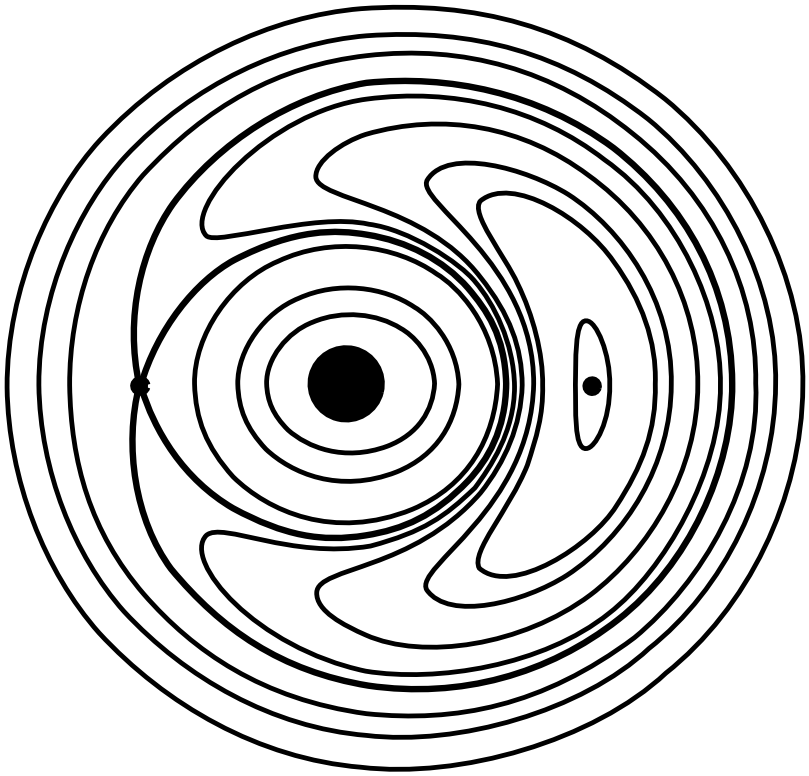


Figura 7.2: El espacio fase del péndulo simple. Se ha mapeado el cilindro sobre un plano con un agujero, se han destacado los dos puntos críticos y se han trazado las separatrices con lineas más gruesas. Podemos ver las tres foliaciones por toros, en este caso toros de una dimensión, separadas por las separatries, se puede distinguir la foliación correspondiente a las libraciones de las correspondientes a las rotaciones.

tomamos a la energía total como integral de movimiento, estas singularidades corresponden a los valores $-mgl$ y mgl de esa integral.

En el caso general tendremos n campos vectoriales independientes V_{I_i}, tangentes a las subvariedades M_c, que conmutan y cuyos flujos sobre esas subvariedades dejan invariantes a las n integrales de movimiento; esto nos hace ver que podemos utilizar a las lineas de flujo de esos n campos como curvas coordenadas sobre M_c. Se puede demostrar que las únicas variedades de dimensión n que admiten esto deben tener la topología de $R^k \times T^{n-k}$ esto es, son cilindros, entendidos como el producto cartesiano de un toro y un espacio cartesiano, en el caso en que M_c sea acotada, entonces necesariamente serán toros de n dimensiones. [3] En el caso de dos dimensiones es claro que se puede trazar una red coordenada sobre un toro, pero no sobre una esfera, sobre la cual llegamos siempre a una singularidad, lo mismo ocurre sobre cualquier otra variedad acotada de dos dimensiones distinta del toro; sobre las no acotadas siempre lo podemos hacer si son difeomorfas a un plano o a un cilindro.

Concluimos así que el espacio fase de un sistema integrable queda estructurado por medio de foliaciones locales compuestas de cilindros de n dimensiones, los que a su vez están generados por haces de trayectorias; estas foliaciones locales quedan separadas por hipersuperficies singulares que llamamos separatrices. En regiones acotadas del espacio fase, esto es regiones en que las trayectorias del sistema no puedan irse al infinito, los cilindros se reducen a toros de n dimensiones. El caso ilustrativo del péndulo es claro al respecto, ya que, por ser de un grado de libertad y toda trayectoria ser acotada, tendremos únicamente toros de una dimensión, esto es curvas cerradas, y, como hicimos ver, dos puntos de equilibrio y separatrices.

En el caso general acotado las trayectorias transcurren sobre toros y, por tanto son el recorrer de n curvas cerradas, o sea, serán la yuxtaposición de n movimientos periódicos, uno por cada una de las n curvas cerradas que generan al toro de n dimensiones. Así, el movimiento será periódico o cuasiperiódico dependiendo de que las n frecuencias sean conmensurables o inconmensurables, esto a su vez dependerá de los valores de las integrales de movimiento y por tanto del toro en el que esté la trayectoria. En resumen en una región acotada del espacio fase de un sistema integrable, las trayectorias quedan organizadas sobre toros de dimensión n que forman foliaciones locales del espacio fase, cada uno de estos toros esta determinado por los valores de las integrales de movimiento, los que también determinan las n frecuencias en que se descomponen las trayectorias sobre él. Como el valor de las frecuencias

[3] Para una demostración de esto ver Arnold...

debe cambiar continuamente con los valores de las integrales de movimiento, la foliación estará compuesta de una sucesión de toros en la que, en general, se intercalan toros con frecuencias conmensurabes entre los que corresponden a frecuencias inconmensurables, esto ocurre de la misma manera en que se intercalan los números racionales entre los irracionales.

Otra forma de ver esto mismo se logra utilizando el teorema de Noether. La existencia de n integrales de movimiento independientes y en involución implica la simetría del sistema mecánico ante un grupo de Lie de n dimensiones cuyas constantes de estructura son todas cero, pues los paréntesis de poisson de las integrales de movimiento se relacionan con los conmutadores de los generadores infinitesimales del grupo. Ahora todo grupo de Lie de n dimensiones con constantes de estructura cero es isomorfo a un grupo de Lie abeliano de n dimensiones. Estos grupos son básicamente R^n bajo la operación de suma o bajo la operación de suma módulo una constante, por ello la variedad que los sustenta debe ser un cilindro de la forma $R\times\cdots\times R\times S^1\times\cdots\times S^1$ donde habrá tantas copias de R como operaciones de suma y tantas de S^1 como operaciones de suma módulo una costante; naturalmente las acciones del grupo sobre el espacio fase generan inmersiones de su variedad en dicho espacio fase y como ésta es un cilindro tendremos que el espacio fase queda foliado por cilindros de dimensión n.

Tomemos al tiro parabólico como ejemplo. En coordenadas cartesianas (x,y) para el plano vertical, la hamiltoniana será

$$H=\frac{1}{2m}\left(p_x^2+p_y^2\right)+mgy, \tag{7.23}$$

el sistema es integrable, pues tanto la energía total como la componente x del impulso lineal son integrales de movimiento y están en involución. Al despejar los impulsos las ecuaciones (7.2) quedan como

$$p_x=c_1, \tag{7.24}$$

$$p_y=\left[2m(c_2-mgy)-c_1^2\right]^{\frac{1}{2}}, \tag{7.25}$$

donde c_1 es el valor de la componente x del impulso lineal y c_2 el de la energía total. La uno forma

$$\omega=c_1dx+\left[2m(c_2-mgy)-c_1^2\right]^{\frac{1}{2}}dy, \tag{7.26}$$

es integrable y al integrarla obtenemos

$$F=c_1x+\int\left[2m(c_2-mgy)-c_1^2\right]^{\frac{1}{2}}dy. \tag{7.27}$$

A partir de (7.14) obtenemos

$$Q^1 = x - \int \frac{c_1}{\left[2m(c_2 - mgy) - c_1^2\right]^{\frac{1}{2}}} dy, \tag{7.28}$$

$$Q^2 = \int \frac{m}{\left[2m(c_2 - mgy) - c_1^2\right]^{\frac{1}{2}}} dy, \tag{7.29}$$

por lo que

$$Q^1 = x + \frac{c_1\sqrt{2mc_2 - c_1^2 - 2m^2gy}}{m^2g} \tag{7.30}$$

y

$$Q^2 = -\frac{\sqrt{2mc_2 - c_1^2 + 2m^2gy}}{mg}. \tag{7.31}$$

La hamiltoniana en el nuevo sistema canónico (Q_1, Q_2, c_1, c_2)es

$$H = c_2 = E, \tag{7.32}$$

por lo que

$$\dot{Q}^1 = 0, \tag{7.33}$$

$$\dot{Q}^2 = 1, \tag{7.34}$$

o

$$Q^1 = \beta^1, \tag{7.35}$$

$$Q^2 = t + \beta^2, \tag{7.36}$$

donde las β son constantes de integración que quedarán determinadas por las condiciones iniciales. A partir de estas ecuaciones podemos obtener la solución en las coordenadas originales utilizando la transformación (7.31). Pero resulta más ilustrativo ver que las trayectorias efectivamente caen sobre subvariedades del espacio fase. Tomemos el espacio fase con coordenadas (x, y, p_x, p_y), una de las integrales de movimiento es p_x, al asignarle un valor c_1 tendremos una subvariedad cartesiana de dimensión tres con coordenadas (x, y, p_y) y en ella una subvariedad dada por $H = c_2$, esto es

$$\frac{1}{2m}\left(c_1^2 + p_y^2\right) + mgy = c_2, \tag{7.37}$$

o

$$p_y^2 = 2m(c_2 - mgy) - c_1^2, \tag{7.38}$$

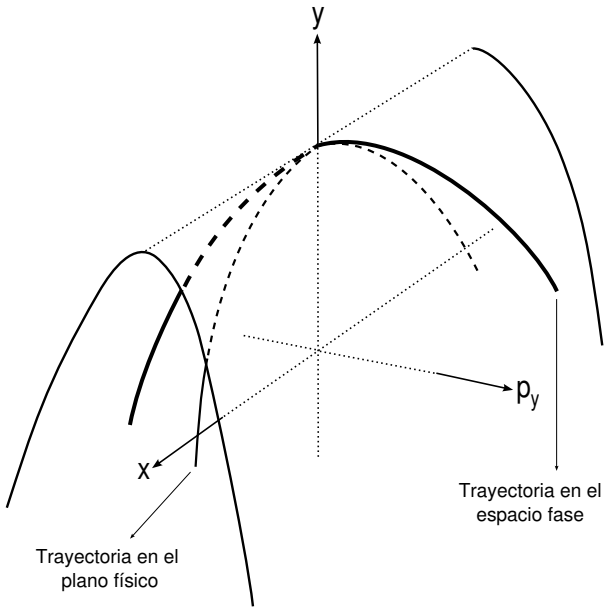


Figura 7.3: El espacio fase del tiro parabólico en el que se muestra una trayectoria sobre un cilindro parabólico invariante y su proyección al plano físico (x, y), se muestra esto desde el subespacio tridimensional $p_x = \text{cst}$.

que es un cilindro parabólico[4] con eje paralelo al eje de las x. Las trayectorias son las curvas que se obtienen al intersectar este cilindro con planos paralelos al eje y y que hacen un ángulo con el eje de las x cuya pendiente es c_1/m, como se muestra en la figura 7.3. Las proyecciones de estas trayectorias sobre el plano xy nos dan las trayectorias parabólicas.

Analicemos el caso del sistema kepleriano. En coordenadas polares la hamiltoniana es

$$H = \frac{1}{2m}\left(p_r^2 + \frac{1}{r^2}p_\theta^2\right) - \frac{k}{r}. \tag{7.39}$$

Tenemos dos integrales de movimiento, que podemos tomar como el impulso angular y la energía, que están en involución, tendríamos entonces

$$p_\theta = l = c_1, \tag{7.40}$$

$$H = E = c_2. \tag{7.41}$$

[4]En este caso no hay regiones acotadas del espacio fase, puesto que no hay trayectorias acotadas.

La primera ecuación nos da directamente p_θ y de la segunda obtenemos

$$p_r = \left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}, \tag{7.42}$$

por tanto tendremos que

$$\omega = \left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}} dr + l d\theta = dF \tag{7.43}$$

y entonces

$$F = l\theta + \int \left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}} dr. \tag{7.44}$$

tenemos entonces que las coordenadas canónicas conjugadas de l y E serán

$$Q^1 = \frac{\partial F}{\partial l} = \theta - \int \frac{l}{r^2\left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}} dr, \tag{7.45}$$

$$Q^2 = \frac{\partial F}{\partial E} = \int \frac{m}{\left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}} dr. \tag{7.46}$$

Las ecuaciones de movimiento y sus soluciones en estas coordenadas son

$$\dot{Q}^1 = 0, \qquad Q^1 = \beta^1, \tag{7.47}$$

$$\dot{Q}^2 = 1, \quad Q^2 = t + \beta^2, \tag{7.48}$$

$$\dot{E} = 0, \qquad E = \text{cst.}, \tag{7.49}$$

$$\dot{l} = 0, \qquad l = \text{cst.} \tag{7.50}$$

por tanto, en las coordenadas originales, tendremos

$$t + \beta^2 = \int \frac{m}{\left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}} dr, \tag{7.51}$$

$$\theta - \beta^1 = \int \frac{l}{r^2\left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}} dr. \tag{7.52}$$

Estas ecuaciones son las que encontramos con anterioridad y el sistema queda reducido a cuadraturas.

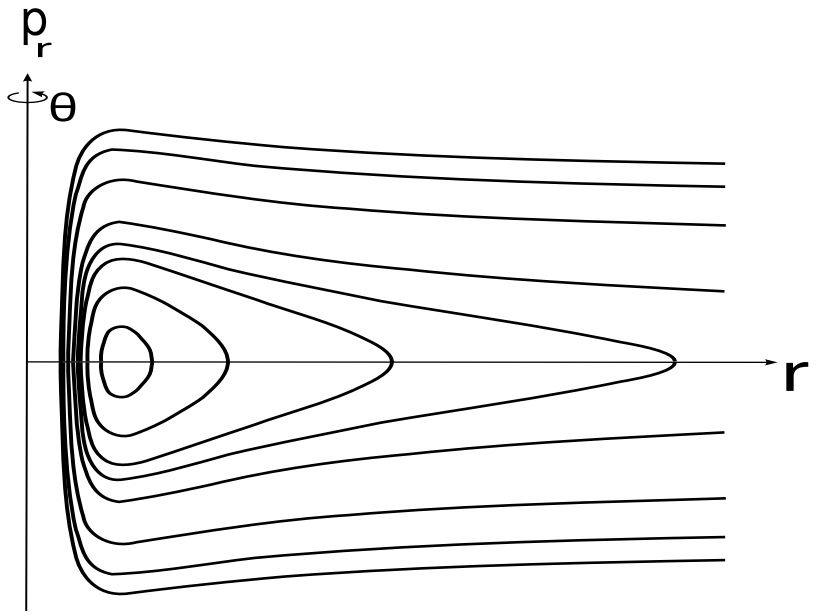


Figura 7.4: En el subespacio p_θ constante del espacio fase del problema de Kepler, las trayectorias cerradas trazarán toros y las abiertas unos cilindros doblados. Esto se aprecia en esta figura al considerar que se debe girar en θ como se indica. Las órbitas parabólicas generan una superficie separatriz entre los toros y los cilindros.

En este caso las trayectorias permanecen acotadas si la energía es negativa y las subvariedades serán toros, en caso de que la energía fuese positiva las trayectorias no quedan acotadas y las subvariedades serán cilindros. Para ver esto tomemos el espacio fase, que en este caso es un espacio cartesiano de cuatro dimensiones y pongamos coordenadas $(r, \theta, p_r, p_\theta)$. Como una de las integrales es p_θ al tomarla como una constante l obtenemos un subespacio R^3. En éste las superficies de energía constante serán las subvariedades sobre la que permanecen las trayectorias, y quedan definidas por

$$\frac{1}{2m}\left(p_r^2 + \frac{1}{r^2}l^2\right) - \frac{k}{r} = E. \tag{7.53}$$

Sobre un plano (p_r, r) tendremos

$$p_r = \pm\left[2m\left(E + \frac{k}{r}\right) - \frac{l^2}{r^2}\right]^{\frac{1}{2}}. \tag{7.54}$$

Para energías negativas esta es la ecuación de una curva cerrada en la que r varía de una $r_{\text{máx}}$ a una $r_{\text{mín}}$ dadas por las raíces de

$$2m\left(E+\frac{k}{r}\right)-\frac{l^2}{r^2}=0. \tag{7.55}$$

Estas curvas se muestran en la figura 7.4 y, como θ es arbitraría en el espacio tridimensional, estas curvas rotan generando toros bisectados por el plano $p_r = 0$. A medida que E aumenta, pero conservándose negativa, se van obteniendo toros concéntricos que constituyen una foliación de la región de energía negativa como se aprecia en la figura; las trayectorias estarán sobre estos toros y, en el caso de Kepler, deben cerrarse pues toda solución es periódica, esto ocurre al ser el tiempo en que varía θ de 0 a 2π igual al que tarda en recorrerse la curva cerrada que genera al toro. La órbita será la proyección de la trayectoria en el toro sobre el plano (r,θ), en este caso es una elipse. A medida que E aumenta, siendo ya positiva, se obtiene, al girar las curvas, una secuencia de cilindros que constituyen una foliación de la región de energía positiva, las trayectorias estarán sobre estos cilindros y al proyectarse sobre el plano (r,θ) generan las órbitas hiperbólicas. Para $E = 0$ tendremos una curva sobre el plano (p_r, r) cuyas ramas superior $(p_r > 0)$ e inferior $(p_r < 0)$ tienden asintóticamente a unirse cuando $r \to \infty$, al hacerla girar sobre θ genera la separatriz entre las dos foliaciones y las proyecciones de las trayectorias sobre ella son las órbitas parabólicas.

En un caso en el que el potencial no sea el de Kepler, pero estemos en una región acotada, será todo similar, excepto que las trayectoria sobre los toros no serán en general cerradas y al proyectar sobre el plano (r,θ) obtenemos las rosetas que ya estudiamos en el capítulo ****. Naturalmente una imagen completa implica variar también el impulso angular, que aquí hemos mantenido fijo, pero para ver la foliación resultante es necesario hacerlo en cuatro dimensiones, aquí tenemos secciones de esa foliación en el subespacio de tres dimensiones correspondiente a $p_r = l$.

En resumen hemos encontrado que todo sistema hamiltoniano integrable es reducible a cuadraturas y su espacio fase queda organizado por el flujo hamiltoniano en regiones foliadas por subvariedades de dimensión n difeomorfas a $T^k \times R^{n-k}$ que llamamos cilindros; estos a su vez están constituidos por haces de trayectorias y están parametrizados por los valores de las integrales de movimiento. Las fronteras entre estas regiones son hipersuperficies singulares de dimensión n que llamamos separatrices. En el caso de encontrarnos en una región acotada del espacio fase, esto es, una en la que las trayectorias permanecen acotadas, la foliación consiste de toros de dimensión n, por lo que toda trayectoria representa un movimiento periódico o cuasiperiódico,

el que sea lo uno o lo otro, depende únicamente del toro sobre el que se desarrolla, esto es del valor de las integrales de movimiento.

7.2. Variables de acción y ángulo.

En una región acotada del espacio fase de un sistema integrable, las trayectorias se desarrollan en el tiempo sobre la superficie de un toro determinado por los valores de las integrales de movimiento, por tanto, son una combinación de n movimientos periódicos. En un sistema arbitrario de coordenadas esto es difícil de observar, pero el hecho de que los toros puedan ser especificados por las integrales de movimiento, nos permite determinar cualquier punto del espacio fase usando como coordenadas a esas integrales de movimiento y a un conjunto de n coordenadas angulares sobre el toro, en esas coordenadas las n periodicidades sería evidentes. Ahora bien unas coordenadas así construidas no son, como sería deseable, necesariamente canónicas; sin embargo, como veremos a continuación, siempre es posible construir sistemas de coordenadas canónicas tales que la mitad de ellas especifiquen los toros y las otras sean coordenadas angulares sobre dicho toros, a esta coordenadas canónicas las llamamos variables de acción y ángulo. Veamos esto.

Consideremos la expresión (7.12) para un sistema integrable en una región acotada del espacio fase. Al dar valores a las integrales de movimiento esta es una diferencial exacta de una función de un toro en los reales, por lo que la integral cíclica

$$\oint \sum_i p_i dq^i \tag{7.56}$$

tomada en una curva cerrada sobre el toro será cero si esta curva es colapsable en un punto sin salirnos del toro, pero será distinta de cero si la curva cerrada no es colapsable en un punto sin salirnos del toro. En un toro se pueden trazar curvas cerradas colapsables y no colapsables como se ilustra en la figura 7.5.

Prestemos atención a la integral (7.56) cuando se toma sobre dos curvas no colapsables γ y γ' que sean deformables una en la otra, resulta ser que el valor de la integral será el mismo sobre las dos curvas. Para ver esto consideremos la figura 7.6 en la que a partir de γ y γ' construimos una nueva curva Γ que es colapsable en un punto, por lo que

$$\oint_\Gamma \sum_i p_i dq^i = 0. \tag{7.57}$$

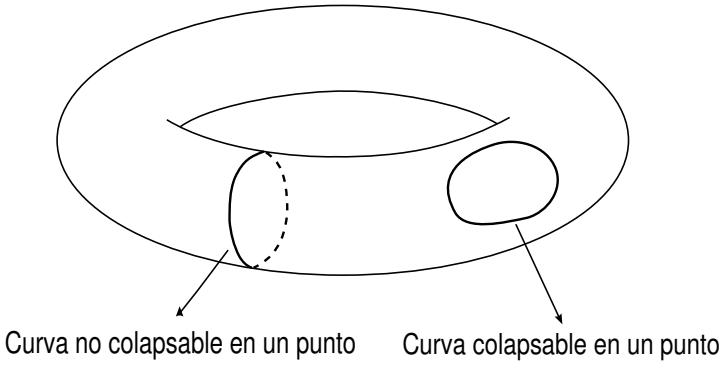


Figura 7.5: Sobre un toro se pueden trazar curvas cerradas deformables o colapsables en un punto y curvas que no pueden deformarse en un punto sin salirse del toro.

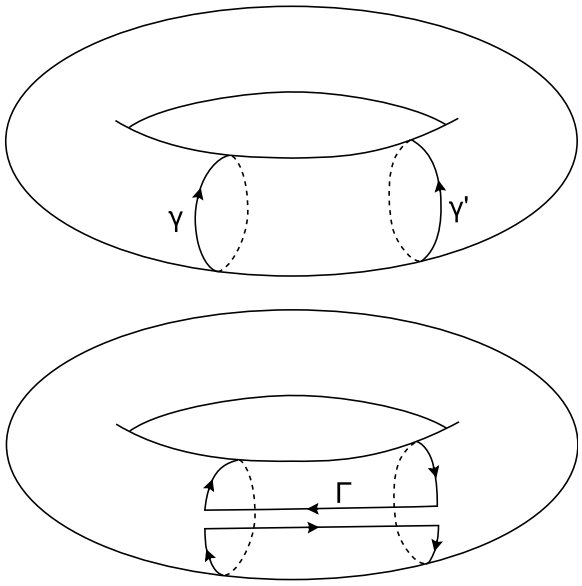


Figura 7.6: Sobre un toro las integrales cíclicas sobre dos curvas no colapsables, pero deformables una en la otra, valen lo mismo.

Esta integral será la suma de las integrales de linea sobre los arcos que componen la curva Γ, por lo que al hacer tender a cero la apertura en γ y γ' tendremos que

$$0 = \oint_\Gamma \sum_i p_i dq^i = \oint_\gamma \sum_i p_i dq^i - \oint_{\gamma'} \sum_i p_i dq^i + \int_{s_1} \sum_i p_i dq^i + \int_{s_2} \sum_i p_i dq^i, \tag{7.58}$$

pero las integrales sobre s_1 y s_2 son integrales tomadas sobre el mismo segmento recorrido en sentidos opuestos, por lo que sus contribuciones se anulan y encontramos que las integrales sobre γ y γ' son iguales. Esto nos permite definir n cantidades que denotaremos por J_j, una por cada una de las n posibles clases de curvas no deformables en un punto sobre el toro de n dimensiones:

$$J_j = \frac{1}{2\pi} \oint_{\gamma_j} \sum_i p_i dq^i, \tag{7.59}$$

donde γ_j es cualquier curva no deformable en un punto representativa de la clase j. A estas cantidades las llamaremos acciones.

Podemos ver de su misma definición que las acciones son funciones únicamente de las integrales de movimiento I, por lo que serán también integrales de movimiento, además son independientes, por ser las curvas γ_j indeformables unas en otras, así podrían constituir una mitad de un sistema de coordenadas; si además estuvieran en involución, entonces ese sistema de coordenadas podría ser canónico. Ahora bien las J son funciones únicamente de las I, que están en involución, por lo que también ellas lo estarán, ya que

$$\{J_i, J_j\} = \sum_{lk} \frac{\partial J_i}{\partial I_l} \frac{\partial J_J}{\partial I_k} \{I_l, I_k\}. \tag{7.60}$$

Hemos así encontrado un nuevo conjunto de n integrales de movimiento independientes y en involución, podemos ahora, siguiendo el procedimiento anteriormente establecido, encontrar un sistema canónico de coordenadas en el que las J sean los impulsos.

Los valores de las coordenadas J, por determinar y estar determinados por las I, fijan, al igual que las I, los toros, veamos como se comportan sus canónicas conjugadas que denotaremos por θ. Tomemos una de ellas θ^i y analicemos su cambio al recorrer una de las curvas no colapsables en un punto γ_j que se utilizaron para encontrar las J. La ecuación (7.14) nos indica que

$$\theta^i = \frac{\partial F}{\partial J_i} \tag{7.61}$$

con

$$F = \int \sum_k p_k dq^k, \tag{7.62}$$

y llamando $\triangle_j \theta^i$ al cambio de θ^i al recorrer la curva γ_j tendremos que

$$\triangle_j \theta^i = \oint_{\gamma_j} d\theta^i = \frac{\partial}{\partial J_i} \oint_{\gamma_j} dF = \frac{\partial}{\partial J_i} \oint_{\gamma_j} \sum_k p_k dq^k \tag{7.63}$$

y recordando (7.59) obtenemos

$$\triangle_j \theta^i = \frac{\partial}{\partial J_i} 2\pi J_j = 2\pi \delta^i_j. \tag{7.64}$$

Vemos que al recorrer una de las curvas cerradas sobre el toro la coordenada θ correspondiente cambia por 2π, mientras que las demás no cambian, esto es las coordenadas θ conforman un sistema de coordenadas angulares sobre el toro, que es precisamente lo que buscábamos.

En resumen, para todo sistema integrable, en una región acotada del espacio fase podemos construir un sistema de coordenadas canónicas en las que los impulsos son integrales de movimiento y determinan a uno de los toros y las coordenadas conjugadas de estos impulsos son ángulos que determinan la posición sobre dicho toro, a estos sistemas de coordenadas se les llama de acción y ángulo, y a las coordenadas en sí, variables de acción y ángulo.

Naturalmente, al expresar la función hamiltoniana en variables de acción y ángulo, ésta resultará función únicamente de las acciones, por lo que las ecuaciones de movimiento serán

$$\dot{\theta}^i = \frac{\partial}{\partial J_i} H(J_1, \ldots, J_n) = \omega^i(J_1, \ldots, J_n), \tag{7.65}$$

$$\dot{J}_i = 0. \tag{7.66}$$

Las ω son constantes por ser las J integrales de movimiento, esto es, podemos decir que a cada toro le corresponden ciertos valores de las ω. Ahora podemos integrar estas ecuaciones y obtener

$$\theta^i = \omega^i t + \beta^i, \tag{7.67}$$

$$J_i = cst., \tag{7.68}$$

lo que nos hace ver que las ω no son otra cosa que las frecuencias angulares del movimiento sobre el toro[5], por lo que, de ser valores conmensurables, el

[5]En caso de haber definido las variables de acción omitiendo el factor $1/2\pi$ en (7.59), obtendríamos las frecuencias en ciclos por segundo, las se acostumbra denotar con la letra ν.

movimiento será periódico o, de ocurrir lo contrario, cuasiperiódico. El poder determinar estas frecuencias para cada uno de los toros, esto de los valores de las J, es una de las ventajas de utilizar las variables de acción y ángulo.

Tomemos como ejemplo el caso del oscilador armónico. La hamiltoniana en coordenadas (x, p) es

$$H = \frac{p^2}{2m} + \frac{k}{2}x^2, \tag{7.69}$$

que es integral de movimiento y, por tratarse de un grado de libertad, el sistema es integrable, su valor será E. Tendremos entonces

$$p = \pm\sqrt{m(2E - kx^2)} \tag{7.70}$$

y

$$J = \frac{1}{2\pi}\oint p dx = \frac{1}{\pi}\int_{-\sqrt{2E/k}}^{\sqrt{2E/k}} \sqrt{m(2E - kx^2)}dx = E\sqrt{\frac{m}{k}}. \tag{7.71}$$

Esto último debido a que la integral cíclica se toma desde el valor mínimo de x hasta el máximo y de retorno, tomando en una dirección el sigo positivo del radical y en la otra el signo negativo. La hamiltoniana en términos de J será

$$H = \sqrt{\frac{k}{m}}J, \tag{7.72}$$

por lo que la frecuencia del movimiento es

$$\omega = \frac{\partial H}{\partial J} = \sqrt{\frac{k}{m}}. \tag{7.73}$$

Ahora en términos de J

$$p = \pm\sqrt{m(2\omega J - kx^2)}, \tag{7.74}$$

por lo que

$$\theta = \frac{\partial}{\partial J}\int p dx = \omega\sqrt{m}\int \frac{dx}{\sqrt{(2\omega J - kx^2)}} = \text{arcsen}\sqrt{\frac{k}{2E}}x. \tag{7.75}$$

En las variables de acción y ángulo las ecuaciones de movimiento son

$$\dot{\theta} = \omega, \tag{7.76}$$

$$\dot{J} = 0 \tag{7.77}$$

con soluciones

$$\theta = \omega t + \beta \tag{7.78}$$

$$J = \text{cst.} \tag{7.79}$$

Al substituir estas soluciones en (7.75) obtenemos

$$x = \sqrt{\frac{2E}{k}}\text{sen}(\omega t + \beta) \tag{7.80}$$

que no es otra que la solución en las coordenadas originales.

Uno de los aspectos más trascendentes de las variables de acción y ángulo lo encontramos en la teoría cuántica de Bohr Somerfeld; en ésta se postula que el comportamiento cuántico de un sistema clásico se obtiene al restringir los valores que pueden adquirir las variables de acción a múltiplos enteros de la constante de Plank sobre 2π. Así para el oscilador armónico tendríamos que

$$J = \frac{E}{\omega} = n\hbar \tag{7.81}$$

y por tanto la energía sólo podría tomar los valores discretos dados por

$$E_n = n\hbar\omega = nh\nu. \tag{7.82}$$

Es importante hacer ver que esta teoría cuántica puede aplicarse únicamente a sistemas que clásicamente son integrables y sólo en regiones acotadas del espacio fase, pues de no ser así el sistema no admite variables de acción ángulo.

En la práctica el cálculo de las variables de acción por medio de (7.59) puede resultar engorroso, sin embargo cuando tenemos un sistema integrable expresado en unas coordenadas canónicas en las que la ecuación de Hamilton Jacobi es totalmente separable, el cálculo se simplifica. En un caso así una solución completa de la ecuación de Hamilton Jacobi se expresa por

$$F(q,\alpha) = W_1(q^1,\alpha) + \cdots + W_n(q^n,\alpha) + \alpha_1 t \tag{7.83}$$

y tendremos que

$$p_i = \frac{\partial W_i(q^i,\alpha)}{\partial q^i} \tag{7.84}$$

y haciendo que la curva cerrada γ_i fuese la intersección del toro correspondiente a los valores de las α con el plano coordenado (q^i, p_i), (7.59) se reduce a

$$J_i = \frac{1}{2\pi}\oint_{\gamma_i} \frac{\partial W_i(q^i,\alpha)}{\partial q^i} dq^i, \tag{7.85}$$

que es mucho más simple de calcular.

Tomemos, para ilustrar este procedimiento, el caso del movimiento planetario y hagámoslo completo en tres dimensiones espaciales. Sabemos que en la región del espacio fase correspondiente a energías negativas, las trayectorias son acotadas y periódicas, ya que las órbitas son elípticas, por lo que en esa región se puede construir un sistema de coordenadas de acción y ángulo.

La función hamiltoniana en coordenadas esféricas es

$$H = \frac{1}{2m}\left[p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\text{sen}^2\theta}p_\varphi^2\right] - \frac{k}{r} \tag{7.86}$$

por lo que la ecuación de Hamilton Jacobi será

$$\frac{1}{2m}\left[\left(\frac{\partial F}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial F}{\partial \theta}\right)^2 + \frac{1}{r^2\text{sen}^2\theta}\left(\frac{\partial F}{\partial \varphi}\right)^2\right] - \frac{k}{r} = -\frac{\partial F}{\partial t}. \tag{7.87}$$

Esta ecuación es separable, lo que hacemos ver proponiendo que

$$F = W_r + W_\theta + W_\varphi - \alpha_1 t \tag{7.88}$$

obtenemos que

$$\frac{1}{2m}\left[\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{1}{r^2\text{sen}^2\theta}\left(\frac{\partial W_\varphi}{\partial \varphi}\right)^2\right] - \frac{k}{r} = \alpha_1. \tag{7.89}$$

Como toda la dependencia en φ está en el término $\partial W_\varphi/\partial\varphi$, éste debe ser constante, esto es

$$p_\varphi = \frac{\partial W_\varphi}{\partial \varphi} = \text{cst.} \tag{7.90}$$

al substituir en (7.89) obtenemos

$$\frac{1}{2m}\left[r^2\left(\frac{\partial W_r}{\partial r}\right)^2 + \left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{p_\varphi^2}{\text{sen}^2\theta}\right] - kr = \alpha_1 r^2. \tag{7.91}$$

Ahora toda la dependencia en θ está en los dos últimos términos dentro del paréntesis, por lo que deben sumar una constante

$$\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{p_\varphi^2}{\text{sen}^2\theta} = \alpha_2^2 \tag{7.92}$$

y por tanto

$$p_\theta = \frac{\partial W_\theta}{\partial \theta} = \left[\alpha_2^2 - \frac{p_\varphi^2}{\text{sen}^2\theta}\right]^{\frac{1}{2}}, \tag{7.93}$$

donde hemos puesto α_2 al cuadrado para indicar que se trata siempre de una cantidad positiva. Finalmente tendremos que

$$\frac{1}{2m}\left[r^2\left(\frac{\partial W_r}{\partial r}\right)^2 + \alpha_2^2\right] - kr = \alpha_1 r^2 \tag{7.94}$$

y

$$p_r = \frac{\partial W_r}{\partial r} = \left[2m(\alpha_1 + \frac{k}{r}) - \frac{\alpha_2^2}{r^2}\right]^{\frac{1}{2}}. \tag{7.95}$$

Hemos así comprobado la separabilidad y encontrado las funciones (7.84).

Pasemos ahora a obtener las variables de acción utilizando (7.85) y recordando que esto es válido únicamente para la región acotada del espacio fase, esto es para energía total negativa o $\alpha_1 < 0$.

Como se ve en la figura 7.7 al transcurrir un ciclo completo del movimiento la φ cambia por 2π, mientras que la p_φ permanece constante, por lo que en el plano coordenado (φ, p_φ) tendremos

$$J_\varphi = \frac{1}{2\pi}\oint p_\varphi d\varphi = p_\varphi. \tag{7.96}$$

Para el caso de J_θ tendremos

$$J_\theta = \frac{1}{2\pi}\oint p_\theta d\theta = \frac{1}{2\pi}\oint \left[\alpha_2^2 - \frac{p_\varphi^2}{\text{sen}^2\theta}\right]^{\frac{1}{2}} d\theta, \tag{7.97}$$

ahora la figura y la ecuación (7.93) nos hacen ver que al transcurrir un ciclo completo θ va de un valor máximo a uno mínimo y regresa al máximo, estos valores estarán dados por los ceros del radical en (7.93), esto es

$$\theta_- = \text{arcsen}\left|\frac{p_\varphi}{\alpha_2}\right|, \tag{7.98}$$

$$\theta_+ = \pi - \theta_- \tag{7.99}$$

al realizar el ciclo p_θ tomará valores en la rama positiva del radical cuando θ vaya en un sentido y en la rama negativa cuando θ vaya en el otro sentido,

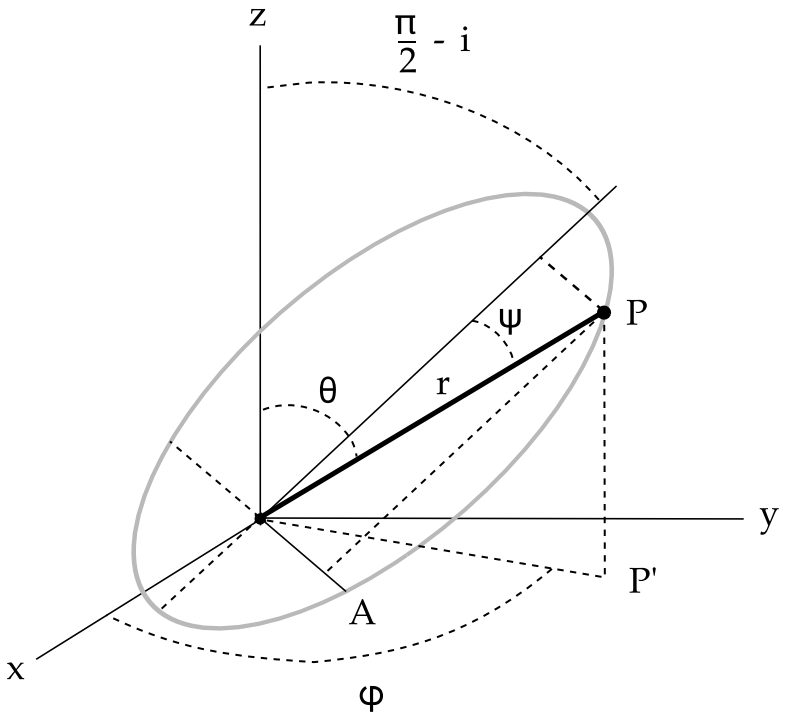


Figura 7.7: El problema de Kepler en el espacio de tres dimensiones. Se ha trazado una órbita para señalar el plano en que se ubica el movimiento. Esta figura se utiliza únicamente para hacer un cambio de variables en la integral para obtener J_θ, en ningún momento se requiere saber que la órbita es una elipse, sólo se requiere saber que permanece sobre un plano.

por lo que la integral cíclica será el doble del valor de la integral definida entre θ_- y θ_+ de ese radical, esto es

$$J_\theta = \frac{1}{\pi}\int_{\theta_-}^{\theta_+}\left[\alpha_2^2 - \frac{p_\varphi^2}{\text{sen}^2\theta}\right]^{\frac{1}{2}} d\theta. \tag{7.100}$$

Esta integral no es sencilla de hacer directamente, sin embargo se puede encontrar J_θ mediante el siguiente procedimiento.

Sabemos que el movimiento ocurre sobre un plano, el de la órbita, si llamamos ψ al ángulo que con respecto al semieje mayor de la órbita hace la linea que une a la partícula con el centro de fuerzas sobre dicho plano, ψ sería la coordenada angular del problema en dos dimensiones cuando se usan coordenadas polares y θ realizará un ciclo cuando ψ cambie por 2π. Por otro lado si llamamos i al ángulo que hace el plano de la órbita con el plano (x, y) tendremos de la figura que $\theta_- = \pi/2 - i$ y $\theta_+ = \pi/2 + i$ y usando 7.98 tendremos

$$p_\varphi^2 = \alpha_2^2 \cos^2 i \tag{7.101}$$

y, de la misma figura, vemos que

$$PP' = r\cos\theta, \tag{7.102}$$

$$PP' = AP\text{sen}i \tag{7.103}$$

y

$$AP = r\cos\psi, \tag{7.104}$$

por lo que

$$\cos\theta = \text{sen}i\cos\psi \tag{7.105}$$

de donde

$$d\theta = \frac{\text{sen}i\text{sen}\psi}{\sqrt{1 - \text{sen}^2 i\cos^2\theta}}d\psi, \tag{7.106}$$

al substituir θ por ψ encontramos

$$J_\theta = \frac{\alpha_2}{\pi}\int_0^\pi \left[\frac{\text{sen}^2 i\text{sen}^2\psi}{1 - \text{sen}^2 i\cos^2\psi}\right] d\psi \tag{7.107}$$

$$= \frac{\alpha_2}{\pi}\int_0^\pi \left[1 - \frac{\cos^2 i}{1 - \text{sen}^2 i\cos^2\psi}\right] d\psi \tag{7.108}$$

$$= \frac{\alpha_2}{\pi}\left\{\pi - \cos i\left[\arctan\frac{\tan\psi}{\cos i}\right]_0^\pi\right\} \tag{7.109}$$

$$= \alpha_2(1 - \cos i) \tag{7.110}$$

$$= \alpha_2 - p_\varphi. \tag{7.111}$$

Podemos observar que

$$\alpha_2 = J_\theta + J_\varphi. \tag{7.112}$$

Pasemos a calcular J_r, de (7.95)

$$J_r = \frac{1}{2\pi} \oint \left[2m(\alpha_1 + \frac{k}{r}) - \frac{\alpha_2^2}{r^2} \right]^{\frac{1}{2}} dr \tag{7.113}$$

$$= \frac{1}{2\pi} \oint \left[2m\alpha_1 r^2 + 2mkr - \alpha_2^2 \right]^{\frac{1}{2}} \frac{dr}{r}. \tag{7.114}$$

Al recorrerse un ciclo del movimiento r irá de $r_{\text{mín}}$ a $r_{\text{máx}}$ y de regreso a $r_{\text{mín}}$, de ida debemos tomar la rama positiva del radical y de regreso la negativa. Esto nos hace ver que, considerando a r como una variable compleja, podemos realizar la integral cíclica sobre la esfera de Riemman usando el contorno γ como se muestra en la figura 7.8. La integral sobre γ será igual a la integral sobre la curva Γ menos la integral sobre una curva δ que encierra al origen, donde hay un polo. Pasemos a calcular las integrales sobre estas curvas.

La integral sobre δ será, por el teorema de Cauchy, $2\pi i$ veces el residuo en el origen que de (7.114) vemos que vale $i\alpha_2$. La integral sobre Γ la podemos hacer considerando que esta curva encierra al punto al infinito, donde hay otro polo. Para calcular el residuo hagamos el cambio de variable

$$u = \frac{1}{r}, \tag{7.115}$$

$$\oint_\Gamma \left[2m\alpha_1 r^2 + 2mkr - \alpha_2^2 \right]^{\frac{1}{2}} \frac{dr}{r} = -\oint_\Gamma \left[2m\alpha_1 + 2mku - \alpha_2^2 u^2 \right]^{\frac{1}{2}} \frac{du}{u^2}, \tag{7.116}$$

multiplicando el integrando por u para calcular el residuo en $u = 0$ tendremos

$$-\left[2m\alpha_1 + 2mku - \alpha_2^2 u^2 \right]^{\frac{1}{2}} \frac{du}{u} \quad = \quad -\left[\sqrt{2m\alpha_1} + \frac{mk}{\sqrt{2m\alpha_1}} u + ... \right] \frac{du}{u}$$

por lo que el residuo será

$$-\frac{mk}{\sqrt{2m\alpha_1}} = -i\sqrt{\frac{mk^2}{-2\alpha_1}}, \tag{7.117}$$

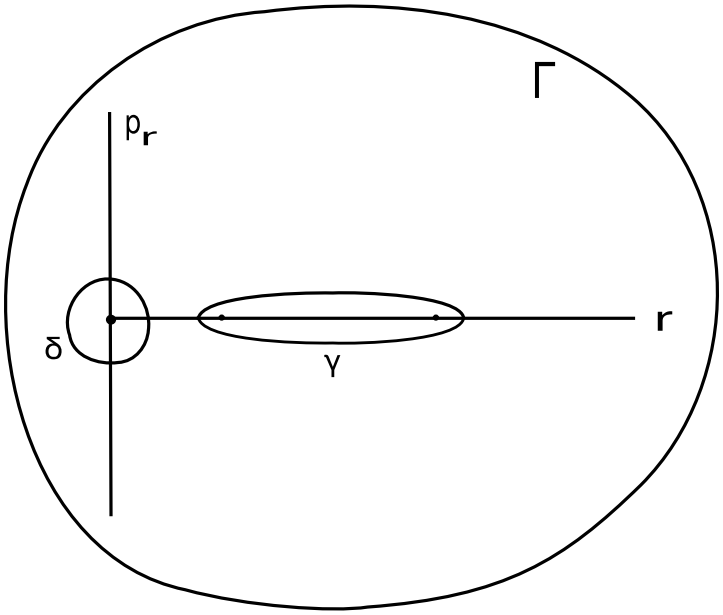


Figura 7.8: Esquema del proceso de integración en la esfera de Riemman para obtener J_r en el problema de Kepler.

encontramos entonces que

$$\oint_{\Gamma} \left[2m\alpha_1 r^2 + kr - \alpha_2^2\right]^{\frac{1}{2}} \frac{dr}{r} = 2\pi\sqrt{\frac{mk^2}{-2\alpha_1}} \tag{7.118}$$

y finalmente al sumar

$$J_r = -\alpha_2 + \sqrt{\frac{mk^2}{-2\alpha_1}}. \tag{7.119}$$

Hemos así calculado las tres variables de acción en función de la la energía total, α_1, del impulso p_φ y de α_2, naturalmente vemos de (7.119) que esto tiene sentido únicamente si $\alpha_1 < 0$.

Ahora de (7.112) tendremos que

$$J_r = -J_\theta - J_\varphi + \sqrt{\frac{mk^2}{-2\alpha_1}}, \tag{7.120}$$

de donde

$$\alpha_1 = E = H = -\frac{mk^2}{2(J_r + J_\theta + J_\varphi)^2}, \tag{7.121}$$

que es la expresión de la hamiltoniana en términos de las variables de acción.

Las frecuencias del movimiento estarán dadas por (7.65); en este caso serán las tres iguales por ser la hamiltoniana una función de la suma de las tres acciones, esto es

$$\omega_r = \omega_\theta = \omega_\varphi = \frac{mk^2}{(J_r + J_\theta + J_\varphi)^3}, \tag{7.122}$$

lo que implica que el movimiento será siempre periódico.

Siempre que la hamiltoniana dependa de la suma de dos variables de acción, las frecuencias correspondientes serán iguales y diremos que, con respecto a esas dos acciones o grados de libertad, el movimiento presenta degeneración. El sistema de Kepler presenta degeneración total, es decir, en todos sus grados de libertad, es por eso que es muy singular ya que en general no se presenta esta degeneración total, de hecho es muy rara. Ahora para los campos centrales (7.112) nos hace ver que siempre se presenta degeneración en J_θ y J_φ, esto es, siempre tendremos dos frecuencias iguales, pero no las tres, cosa que solo ocurre para el campo kepleriano y para el caso $V = r^2$.

En la mecánica cuántica de Bohr Somerfeld debemos substituir las variables de acción por múltiplos enteros de $\hbar$. En el caso del átomo de hidrógeno debemos hacerlo en (7.121), por lo que obtendremos el conocido fenómeno de degeneración de los niveles de energía, pues la energía dependerá únicamente de la suma de los tres números cuánticos. Vemos que, en el contexto de esta teoría cuántica, la degeneración del átomo de hidrógeno es una consecuencia directa de la degeneración clásica del sistema kepleriano.

7.3. Invariantes adiabáticos.

Hay ocasiones en que un sistema mecánico contiene algunos parámetros que varían lentamente con el tiempo, desde luego esto hace que la hamiltoniana del sistema sea dependiente del tiempo y no puede ser ser una integral de movimiento. Si un sistema así fuera integrable para valores constantes de dichos parámetros, cuando éstos cambian con el tiempo el sistema deja de ser integrable y las que eran sus integrales, por ejemplo la energía total, cambiarán significativamente con el tiempo al ir cambiando el parámetro. En ocasiones ocurre que, debido a la lentitud del cambio, alguna variable del sistema, esto es, una función del espacio fase y de los parámetros, permanece aproximadamente constante, a pesar de que, si la variación de los parámetros no fuese lenta, cambiaría significativamente; a una variable así se le llama un invariante adiabático. El término adiabático se refiere al hecho

de que la variación temporal de los parámetros es muy lenta y el término invariante no debe interpretarse como constante, la variable si cambia de valor con el tiempo, en este sentido no permanece invariante, pero su cambio es muy pequeño en comparación con el cambio que sufren los parámetros o la hamiltoniana misma.

Las variables de acción de un sistema que fuese integrable y acotado si sus parámetros permaneciesen constantes, son, con mucho, los más importantes ejemplos de invariantes adiabáticos cuando algún parámetro del sistema varía lentamente con el tiempo. Veamos esto para sistema de un grado de libertad.

Consideremos un sistema de un grado de libertad en una región acotada del espacio fase; pensemos en uno de los parámetros del sistema (puede ser la masa de un partícula, la constante de un resorte o cualquier otro elemento del sistema) y denotémosle con λ. La hamiltoniana en algunas coordenadas canónicas (q,p) será $H(q,p,\lambda)$. Cuando el parámetro es constante las trayectorias serán periódicas y existirán variables de acción y ángulo (θ, J) que naturalmente serán distintas para distintos valores del parámetro. Denotemos por $F(q,J,\lambda)$ a la función generadora de transformaciones canónicas que, dado un valor fijo de λ, transforma las coordenadas (q,p) en las de acción y ángulo (θ, J).

Consideremos ahora que λ cambia con el tiempo, esto es, que es una función del tiempo $\lambda(t)$, pero que lo hace monótona y lentamente, esto es, $\dot{\lambda}$ no cambia de signo y para un tiempo τ característico del sistema

$$|\bigtriangledown\lambda| \approx \left|\dot{\lambda}\tau\right| = \left|\dot{\lambda}\right| \frac{2\pi}{\nu} \ll |\lambda|\,, \tag{7.123}$$

donde el tiempo τ puede ser el período típico del sistema y ν la frecuencia correspondiente.

La hamiltoniana es ahora función del tiempo vía λ y deja de ser una integral de movimiento, naturalmente, el sistema deja de ser integrable. A pesar de esto podemos seguir estudiandolo en el sistema de coordenadas canónicas generadas por la función $F(q,J,\lambda(t))$, la que, por ser ahora dependiente del tiempo, nos obliga a considerar una nueva hamiltoniana

$$H'(\theta,J,\lambda) = H(J,\lambda) + \frac{\partial F}{\partial t} \tag{7.124}$$

$$= H(J,\lambda) + \dot{\lambda}\frac{\partial F(\theta,J,\lambda)}{\partial \lambda}, \tag{7.125}$$

donde en el último término hemos ya expresado el resultado en función de las coordenadas θ y J. La coordenada θ seguirá siendo una coordenada angular, por lo que, F debe ser una función periódica de θ de período 2π.

Las ecuaciones de movimiento serán

$$\dot{\theta} = \frac{\partial H}{\partial J} + \dot{\lambda}\frac{\partial^2 F}{\partial J \partial \lambda}, \tag{7.126}$$

$$\dot{J} = -\dot{\lambda}\frac{\partial^2 F}{\partial \theta \partial \lambda}. \tag{7.127}$$

Como era de esperar J ya no es constante y θ ya no varía linealmente con el tiempo.

El parámetro λ debe cumplir con (7.123), por lo que podemos ponerlo como

$$\lambda(t) = \lambda_0 + \epsilon t + o(\epsilon^2) \tag{7.128}$$

con

$$|\epsilon| \ll \frac{|\lambda|}{\tau_0} = \frac{|\lambda|}{2\pi}\nu_0 = \frac{|\lambda|}{2\pi}\frac{\partial H(J,\lambda_0)}{\partial J} \tag{7.129}$$

donde τ_0 es el período de una oscilación en el caso en que λ fuese constante e igual a λ_0, el valor de λ al tiempo cero, y ν_0 la frecuencia correspondiente. Bajo estas condiciones $\dot{\lambda}$ es del orden $|\epsilon|$ y las ecuaciones de movimiento serán

$$\dot{\theta} = \nu_0 + \epsilon\left(\left.\frac{\partial^2 H}{\partial\lambda\partial J}\right|_{t=0} t + \frac{\partial^2 F}{\partial J\partial\lambda}\right) + o(\epsilon^2), \tag{7.130}$$

$$\dot{J} = -\epsilon\frac{\partial^2 F}{\partial\theta\partial\lambda} + o(\epsilon^2). \tag{7.131}$$

La primera ecuación nos indica que por un tiempo largo θ será monótona en el tiempo y se modificará por 2π cada vez que transcurra un tiempo del orden de $\tau = \tau_0 + o(\epsilon)$.

Calculemos el incremento de J en el tiempo. F debe ser periódica de período 2π en θ, por lo que también lo serán sus derivadas en (7.131), entonces su aportación al incremento de J en el tiempo τ será de orden ϵ, por lo que el incremento de J será de orden ϵ^2, esto es, al transcurrir el tiempo J fluctúa a orden ϵ en torno a un valor constante por tanto, J permanece constante a orden ϵ hasta haber transcurrido un tiempo del orden de magnitud de $1/\epsilon$, que como el cambio es adiabático corresponde a tiempos que abarcan muchos periodos τ_0; vemos que J es un invariante adiabático, esto es, permanece aproximadamente constante durante un tiempo tan largo como queramos si ϵ es suficientemente pequeña. Notemos que otras variables o cantidades importantes del sistema varían a orden ϵ como el propio parámetro λ o la energía, por lo que al transcurrir un tiempo del orden de $1/\epsilon$ habrán cambiado considerablemente, mientras que J sólo lo habrá hecho por una cantidad del orden de ϵ.

Tomemos como ejemplo el oscilador armónico. La hamiltoniana de este sistema es

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 x^2 \tag{7.132}$$

donde $\omega^2 = k/m$ es la frecuencia angular de la oscilación. (7.71) y (7.75) nos indican que las coordenadas de acción y ángulo son

$$J = \frac{H}{\omega} \tag{7.133}$$

y

$$\theta = \text{arcseno}\sqrt{\frac{\omega m}{2J}}x. \tag{7.134}$$

La función generadora de transformaciones canónicas que proporciona estas variables es

$$F(x, J) = \int \sqrt{2m\omega J - \omega^2 m^2 x^2}dx \tag{7.135}$$

$$= J\text{arcseno}\sqrt{\frac{\omega m}{2J}}x + x\frac{\sqrt{2m\omega J - \omega^2 m^2 x^2}}{2}. \tag{7.136}$$

En estas coordenadas la solución del oscilador armónico es

$$\theta = \omega t, \tag{7.137}$$

$$J = \text{cst}. \tag{7.138}$$

donde por comodidad hemos considerado que θ es cero al tiempo cero.

Si ω, o, lo que es lo mismo, la constante del resorte, cambia con el tiempo, el sistema dejará de ser integrable y la función generadora (7.136) será una función del tiempo, en este caso, para tratar el sistema en las coordenadas generadas por la función (7.136) debemos considerar una nueva hamiltoniana dada por

$$H' = H + \frac{\partial F}{\partial t} \tag{7.139}$$

y, como toda la dependencia temporal esta en la frecuencia ω, podemos tener

$$H' = H + \dot{\omega}\frac{\partial F}{\partial \omega}. \tag{7.140}$$

De (7.136) tendremos que

$$\frac{\partial F}{\partial \omega} = \frac{1}{2\omega}x\sqrt{2m\omega J - \omega^2 m^2 x^2}, \tag{7.141}$$

esta cantidad al expresarla en las coordenadas (θ, J) es

$$\frac{\partial F}{\partial \omega} = \frac{J}{2\omega}\text{sen}2\theta. \tag{7.142}$$

Podemos ahora expresar la hamiltoniana H' en las coordenadas (θ, J)

$$H' = \omega J + \dot{\omega}\frac{J}{2\omega}\text{sen}2\theta. \tag{7.143}$$

Las ecuaciones de movimiento serán

$$\dot{\theta} = \omega + \frac{\dot{\omega}}{2\omega}\text{sen}2\theta, \tag{7.144}$$

$$\dot{J} = -\frac{\dot{\omega}}{\omega}\text{cos}2\theta. \tag{7.145}$$

Vemos que en caso de que ω no dependiese del tiempo estas ecuaciones, como era de esperarse, se reducen a las del oscilador armónico.

Consideremos ahora una dependencia temporal de ω que cumpla con las condiciones (7.123), por lo que podemos a primer orden escribir

$$\omega = \omega_0 + \epsilon t, \tag{7.146}$$

con

$$|\epsilon| \ll \frac{\omega_0^2}{2\pi}. \tag{7.147}$$

Tendremos entonces que

$$\frac{1}{\omega} = \frac{1}{\omega_0 + \epsilon t} = \frac{1}{\omega_0} - \frac{t}{\omega_0^2}\epsilon + o(\epsilon^2) \tag{7.148}$$

y

$$\frac{\dot{\omega}}{\omega} = \frac{\epsilon}{\omega_0} + o(\epsilon^2). \tag{7.149}$$

Las ecuaciones de movimiento serán

$$\dot{\theta} = \omega_0 + \epsilon\left(t + \frac{\text{sen}2\theta}{2\omega_0}\right) + o(\epsilon^2), \tag{7.150}$$

$$\dot{J} = -\frac{\epsilon}{\omega_0}J\text{cos}2\theta + o(\epsilon^2). \tag{7.151}$$

Vemos que la variación de $\dot{\theta}$ será pequeña en comparación con su valor inicial ω_0, por lo que esperamos que θ tenga una evolución temporal monótona y cercana a la que tendría si la frecuencia no dependiese del tiempo,

esto nos permite, en la segunda ecuación, substituir la derivada temporal de J por su derivada con respecto a θ, esto es

$$\frac{dJ}{d\theta} = -\frac{\frac{\epsilon}{\omega_0} J\cos 2\theta}{\omega_0 + \epsilon\left(t + \frac{\text{sen}2\theta}{2\omega_0}\right)} + o(\epsilon^2) \tag{7.152}$$

y desarrollando en ϵ tendremos

$$\frac{dJ}{d\theta} = -\frac{\epsilon}{\omega_0^2} J\cos 2\theta + o(\epsilon^2). \tag{7.153}$$

De esta ecuación obtenemos que

$$\frac{dJ}{J} = -\frac{\epsilon}{\omega_0^2}\cos 2\theta d\theta + o(\epsilon^2) \tag{7.154}$$

e integrando entre el tiempo inicial $t = 0$ y un cierto tiempo final $t = t_f$ tendremos que

$$\ln\frac{J_f}{J_0} = -\frac{\epsilon}{\omega_0^2}\int_0^{\theta_f} \cos 2\theta d\theta + o(\hat{e^2}), \tag{7.155}$$

como el coseno es una función periódica su integral estará comprendida entre dos y menos dos por lo que a primer orden en ϵ

$$\exp\left(-\frac{2\epsilon}{\omega_0^2}\right) \leq \frac{J_f}{J_0} \leq \exp\left(\frac{2\epsilon}{\omega_0^2}\right), \tag{7.156}$$

pero (7.147) nos indica que

$$\left|\frac{2\epsilon}{\omega_0^2}\right| \ll \frac{1}{\pi} \tag{7.157}$$

por lo que J_f/J_0 será muy cercano a uno, esto es, a orden ϵ, J es prácticamente constante, presentará únicamente pequeñas fluctuaciones en torno a su valor inicial. Hemos así hecho ver que la variable de acción en un invariante adiabático.

Capítulo 8

Sistemas mecánicos no integrables.

En el capítulo anterior encontramos que todo sistema mecánico integrable es reducible a cuadraturas y establecimos que las trayectorias en el espacio fase se organizan de manera muy regular, generando foliaciones en cilindros o toros. En el caso de las regiones acotadas las soluciones son siempre periódicas o cuasiperiódicas y podemos encontrar sistemas de coordenadas canónicas de acción y ángulo. Las acciones, son integrales de movimiento y sus valores especifican a uno de los toros entre los que conforman la foliación; sus conjugadas, las de ángulo, son coordenadas angulares sobre los toros. Así podemos decir que conocemos el comportamiento de las soluciones de cualquier sistema integrable y que esto agota su estudio general.

Para el caso de los sistemas no integrables, o sea los que no tienen tantas integrales de movimiento independientes y en involución como grados de libertad, no tenemos algo similar. Desde luego no presentarán el comportamiento general de los integrables, pero podrían presentar comportamientos estructurados que de alguna forma nos permitieran conocer las propiedades cualitativas de sus trayectorias y así clasificarlos; de ser esto posible tendríamos un panorama general y completo de los sistemas mecánicos. En este capítulo abordaremos el estudio de los sistemas no integrables, no sin antes señalar que, contrario a lo que ocurre con los integrables, las trayectorias se desarrollan en el espacio fase en formas sumamente complejas por lo que no se ha llegado a un estudio completo.

8.1. Existencia de los sistemas no integrables.

El tratar los sistemas no integrables en lo general resulta sumamente complejo. Es posible tratar los integrables porque tenemos una propiedad muy específica, la existencia de las n integrales de movimiento en involución, a partir de la cual podemos hacer deducciones e iniciar el estudio. Cuando partimos de que el sistema no tiene dicha propiedad, resulta muy difícil siquiera comenzar a hacer deducciones, pues al quedar definidos por una 'no propiedad' se hacen inasibles. En adición a esto se presenta otra dificultad, no se tienen ejemplos de sistema no integrable que haya sido resueltos, esto es que hayan sido reducidos a cuadraturas, la ausencia de ejemplos que inspiren el desarrollo teórico es una dificultad muy seria para avanzar en su estudio.

De hecho debemos comenzar con una pregunta básica, ¿existen los sistemas no integrables?. Esto es, sabemos, ya que tenemos ejemplos de ello, que existen sistemas integrables, pero esto no implica que existan los no integrables. Pudiera ser que aquellos sospechosos de serlo, por ejemplo el problema de tres cuerpos, resultara que si tuvieran dichas integrales, pero que no hemos sido capaces de encontrarlas. Poincaré atacó este problema, precisamente teniendo en mente el problema de tres cuerpos, y llegó a demostrar la existencia de sistemas no integrables, es más llegó a demostrar que conforman una absoluta mayoría de los sistemas mecánicos. Veamos esto con más detenimiento.

Comencemos considerando sistemas cercanos a uno integrable. Esto se hace así por las razones ya enunciadas y porque en muchas situaciones prácticas se presentan sistemas de esta naturaleza. Por ejemplo, al estudiar la influencia que en el movimiento de un planeta ejerce otro, sabemos que la magnitud de la fuerza que ejerce el Sol sobre dicho planeta es mucho mayor que la magnitud de la que ejerce sobre él el otro planeta, esto permite considerar a la influencia del otro planeta como una pequeña adición o perturbación al sistema kepleriano integrable consistente del Sol y el planeta en estudio. Así surge la llamada teoría de perturbaciones que intenta establecer resultados o realizar cálculos sobre sistemas cercanos a un sistema integrable cuya solución es conocida. En términos generales se trata de presentar la solución del sistema perturbado como una serie de potencias en la magnitud de la perturbación que se considera pequeña; el primer término de esta serie, que no depende de la perturbación, es la solución del sistema integrable ya conocida.

En términos hamiltonianos esto se presenta de la siguiente manera. Se tiene, en un espacio fase, un sistema integrable con función hamiltoniana H_0, la perturbación se presenta como la adición a esta hamiltoniana de términos

pequeños para obtener una nueva hamiltoniana

$$H = H_0 + \epsilon H_1 + \epsilon^2 H_2 + \cdots , \tag{8.1}$$

donde ϵ es un parámetro pequeño que regula la magnitud de la perturbación. Nótese que al adicionar la perturbación no se modifican ni el espacio fase ni la forma simpléctica, únicamente cambia la función hamiltoniana. La teoría de perturbaciones parte para el estudio del sistema completo de la solución del sistema integrable H_0 y busca soluciones para H en series de potencias de ϵ. Naturalmente estamos suponiendo que las funciones H_i son diferenciables y que ϵ es suficientemente pequeña como para que la serie sea convergente.

Veamos en primer término el asunto de la integrabilidad del sistema H. Para ello, con objeto de ser lo más claro posible, consideraremos que el sistema tiene únicamente dos grados de libertad, que nos encontramos en una región acotada del espacio fase para la hamiltoniana integrable H_0 y que esta es no degenerada, esto es, en variables de acción y ángulo $(J_1, J_2, \theta_1, \theta_2)$ para H_0, tendremos que:

1. H_0 es función únicamente de J_1 y J_2.
2. Las H_i son funciones periódicas de θ_1 y θ_2 de período 2π.
3. La Hessiana $\frac{\partial H_0(J_1,J_2)}{\partial(J_1,J_2)}$ es distinta de cero.

Bajo estas hipótesis, y para ϵ suficientemente pequeña, en general el sistema hamiltoniano H no tiene integrales de movimiento independientes de H misma. Esto es, la única integral de movimiento es la hamiltoniana y por tanto el sistema es no integrable. Pasemos a demostrar esto siguiendo a Poincaré y Whittaker.

Una integral de movimiento del sistema H será una función $I(J_1, J_2, \theta^1, \theta^2, \epsilon)$ diferenciable, bien definida en la región de definición de las variables acción y ángulo, univaluada y periódica en los ángulos θ con períodos 2π. Entonces para ϵ suficientemente pequeña podríamos escribir

$$I = I_0(J, \theta) + \epsilon I_1(J, \theta) + \epsilon^2 I_2(J, \theta) + \cdots , \tag{8.2}$$

donde las I_j cumplirían las mismas condiciones de diferenciabilidad y periodicidad que I. Supongamos, por reducción al absurdo, que existe para el sistema con hamiltoniana H una integral de movimiento I independiente de H.

Como hemos supuesto que I es una integral de movimiento

$$\{H, I\} = 0, \tag{8.3}$$

por lo que de (8.1) y (8.2) obtenemos

$$\begin{aligned} \{H_0, I_0\} + \epsilon \left[\{H_0, I_1\} + \{H_1, I_0\}\right] + \\ \epsilon^2 \left[\{H_0, I_2\} + \{H_1, I_1\} + \{H_2, I_0\}\right] + \cdots \quad &= \quad 0, \end{aligned}$$

por tanto se debe cumplir que

$$\{H_0, I_0\} = 0, \tag{8.4}$$

$$\{H_0, I_1\} + \{H_1, I_0\} = 0, \tag{8.5}$$

$$\{H_0, I_2\} + \{H_1, I_1\} + \{H_2, I_0\} = 0, \tag{8.6}$$

$$\vdots \tag{8.7}$$

lo que nos hace ver que I_0 es una integral de movimiento del sistema integrable H_0.

Demostraremos ahora que si I es una integral independiente de H,I_0 no puede ser función de H_0, esto es, que I_0 es una integral de movimiento del sistema no perturbado independiente de H_0 o que, de no serlo, existe otra integral de movimiento de H independiente de H cuyo término independiente de ϵ no es función de H_0. Supongamos que I_0 si es dependiente de H_0. Del hecho de que H_0 es función únicamente de las J con Hessiana distinta de cero, podemos despejar, digamos J_1, y obtener

$$J_1 = G(H_0, J_2), \tag{8.8}$$

con G una función univaluada de sus argumentos; entonces remplazando en I_0 tendremos que

$$I_0(J_1, J_2, \theta^1, \theta^2) = F(H_0, J_2, \theta^1, \theta^2), \tag{8.9}$$

pero hemos supuesto que I_0 es función únicamente de H_0, por tanto

$$I_0 = F(H_0), \tag{8.10}$$

donde $F(H_0)$ no es multivaluada.

Al ser F una función diferenciable y univaluada, $F(H)$ será integral de movimiento de H y también lo será $I - F(H)$, por serlo I; ahora ((8.10)) nos muestra que la integral $I - F(H)$ no tiene un término independiente de ϵ, esto es podemos escribir

$$I - F(H) = \epsilon I', \tag{8.11}$$

por lo que hemos encontrado una nueva integral de movimiento I' que podemos desarrollar como

$$I' = I'_0 + \epsilon I'_1 + \epsilon^2 I'_2 + \cdots . \tag{8.12}$$

En general I_0' no será dependiente de H_0, pero si lo fuera, podríamos repetir el procedimiento ahora para I', con lo que obtendríamos una integral de movimiento I'', cuya parte independiente de ϵ no sería en general dependiente de H_0, pero si lo fuera, podríamos continuar con el proceso hasta que ocurriera alguna de dos únicas posibilidades: o bien el proceso termina al obtenerse una integral de movimiento de H cuya parte independiente de ϵ no es una función de H_0 o el proceso no termina nunca, pero en ese caso I resultaría ser una función de H contrario a lo que hemos supuesto. Por tanto, si existe una integral independiente de H, siempre podemos construir una integral de movimiento independiente de H cuya parte independiente de ϵ no es función de H_0.

Llamemos I a esa integral no dependiente de H y cuya parte independiente de ϵ, I_0, no es función de H_0. Demostraremos ahora que I_0 no es función de las coordenadas angulares.

Supongamos que si lo es, entonces debe ser periódica en ellas de período 2π y por tanto podemos escribir

$$I_0 = \sum_{m_1,m_2} A_{m_1,m_2}(J_1,J_2)e^{i(m_1\theta_1+m_2\theta_2)} \tag{8.13}$$

y como H_0 sólo depende de las J tendremos que

$$\begin{aligned}\{I_0,H_0\} &= \frac{\partial I_0}{\partial\theta_1}\frac{\partial H_0}{\partial J_1}+\frac{\partial I_0}{\partial\theta_2}\frac{\partial H_0}{\partial J_2}\\ &= \sum_{m_1,m_2}\left(m_1\frac{\partial H_0}{\partial J_1}+m_2\frac{\partial H_0}{\partial J_2}\right)iA_{m_1,m_2}(J_1,J_2)e^{i(m_1\theta_1+m_2\theta_2)}\\ &= \sum_{m_1,m_2}(m_1\omega_1+m_2\omega_2)\,iA_{m_1,m_2}(J_1,J_2)e^{i(m_1\theta_1+m_2\theta_2)},\end{aligned}$$

donde m_1 y m_2 son todos los enteros positivos y negativos y ω_1 y ω_2 son las frecuencias correspondientes a J_1 y J_2 del sistema integrable H_0. Pero este paréntesis de Poisson es cero por lo que

$$(m_1\omega_1+m_2\omega_2)\,A_{m_1,m_2}(J_1,J_2)=0 \quad \forall \quad m_1,m_2,J_1 \;\; y \;\; J_2, \tag{8.14}$$

entonces $m_1\omega_1+m_2\omega_2=0$ o $A_{m_1,m_2}(J_1,J_2)=0$, el primer caso, excepto cuando $m_1=m_2=0$, implicaría que la hessiana de H_0 fuese cero en contra de nuestra hipótesis, por lo que

$$A_{m_1,m_2}(J_1,J_2)=0 \quad \forall \quad m_1,m_2\neq 0 \tag{8.15}$$

y entonces

$$I_0 = A_{0,0}(J_1,J_2), \tag{8.16}$$

que es lo que queríamos hacer ver.

Tomemos ahora la segunda de las condiciones ((8.5))

$$\{H_0, I_1\} + \{H_1, I_0\} = 0 \tag{8.17}$$

en la que sabemos que I_1 y H_1 son funciones periódicas de θ_1 y θ_2 de períodos 2π y por tanto podemos hacer

$$H_1 = \sum_{m_1,m_2} B_{m_1,m_2}(J_1, J_2) e^{i(m_1\theta_1 + m_2\theta_2)}, \tag{8.18}$$

$$I_1 = \sum_{m_1,m_2} C_{m_1,m_2}(J_1, J_2) e^{i(m_1\theta_1 + m_2\theta_2)}. \tag{8.19}$$

Entonces la condición ((8.5)) nos dice que

$$\begin{aligned} \sum_{m_1,m_2} \Bigg[-m_1 C_{m_1,m_2} \frac{\partial H_0}{\partial J_1} - m_2 C_{m_1,m_2} \frac{\partial H_0}{\partial J_2} \quad + \\ m_1 B_{m_1,m_2} \frac{\partial I_0}{\partial J_1} + m_2 B_{m_1,m_2} \frac{\partial I_0}{\partial J_2} \Bigg] i e^{i(m_1\theta_1 + m_2\theta_2)} \quad = \quad 0, \end{aligned}$$

de donde

$$-C_{m_1,m_2}\left(m_1\frac{\partial H_0}{\partial J_1} + m_2\frac{\partial H_0}{\partial J_2}\right) + B_{m_1,m_2}\left(m_1\frac{\partial I_0}{\partial J_1} + m_2\frac{\partial I_0}{\partial J_2}\right) = 0 \ \ \forall \ \ m_1, m_2, J_1 \ \ y \ \ J_2. \tag{8.20}$$

Para aquellos valores de las J que corresponden a toros resonantes existen valores de las m tales que

$$m_1\frac{\partial H_0}{\partial J_1} + m_2\frac{\partial H_0}{\partial J_2} = 0, \tag{8.21}$$

para estos valores de las J tendremos que

$$B_{m_1,m_2} = 0 \quad o \quad m_1\frac{\partial I_0}{\partial J_1} + m_2\frac{\partial I_0}{\partial J_2} = 0, \tag{8.22}$$

pero las B están determinadas por H_1, por lo que, en general, no serán cero para esos valores de las J y las m que están determinados por la H_0, esto implica que, en general, para los valores en que ((8.21)) se cumple será valido que

$$m_1\frac{\partial I_0}{\partial J_1} + m_2\frac{\partial I_0}{\partial J_2} = 0 \tag{8.23}$$

y entonces

$$\frac{\partial(H_0, I_0)}{\partial(J_1, J_2)} = \begin{vmatrix} \frac{\partial H_0}{\partial J_1} & \frac{\partial H_0}{\partial J_2} \\ \frac{\partial I_0}{\partial J_1} & \frac{\partial I_0}{\partial J_2} \end{vmatrix} = 0 \tag{8.24}$$

sobre los toros conmensurables; pero, dentro del conjunto de todos los toros o valores de las J, los conmensurables forman un conjunto denso, por lo que H_0 e I_0 no son funciones independientes, lo que contradice lo ya dicho y entonces la hipótesis de que I es independiente de H nos lleva a una contradicción, por lo que hemos demostrado que no existe una segunda integral de movimiento independiente de la hamiltoniana y, por tanto, que el sistema perturbado es no integrable. Esta demostración puede ser generalizada a más grados de libertad e incluir algunos casos excepcionales como aquellos en los que, por casualidad, los coeficientes B_{m_1,m_2} determinados por H_1 sean cero sobre los toros conmensurables de H_0.

En realidad hemos demostrado mucho más que la existencia de sistemas hamiltonianos no integrables, hemos demostrado que dado un sistema integrable existe, en el espacio de todas las funciones hamiltonianas, una vecindad de dicho sistema en donde la no integrabilidad es una propiedad genérica y esto se puede extender al conjunto de todos los sistemas hamiltonianos, esto es, los sistemas hamiltonianos son genéricamente no integrales, es más, se puede hacer ver que genéricamente únicamente tienen una integral de movimiento, la hamiltoniana misma.

8.2. Las trayectorias en los sistemas no integrables.

En la sección anterior llegamos a la conclusión de que los sistemas no integrables existen y que, en general, un sistema cercano a uno integrable no será integrable; pero esto no resuelve la pregunta sobre el comportamiento de las trayectorias en los sistemas no integrables. El mismo Poincaré desarrolló un primer acercamiento a este asunto, demostró que en general al perturbar un sistema hamiltoniano integrable se rompen los toros conmensurables, esto es la mayor parte de las soluciones periódicas desaparecen, pero no encontró respuesta a la pregunta sobre el destino de los toros inconmensurables y con ellos de las soluciones cuasiperiódicas. Veamos este importante resultado.

En un sistema integrable acotado H_0 las trayectorias se agrupan formando una estructura de toros de dimensión n que folia cierta región del espacio fase. Al actuar una perturbación el sistema $H = H_0 + \epsilon H_1 + \cdots$ deja, en general, de ser integrable. Pudiera ser que únicamente se deformarán los toros un poco y que la foliación permaneciera como tal y, con ella, el carácter periódico o cuasiperiódico de las trayectorias; o por el contrario,

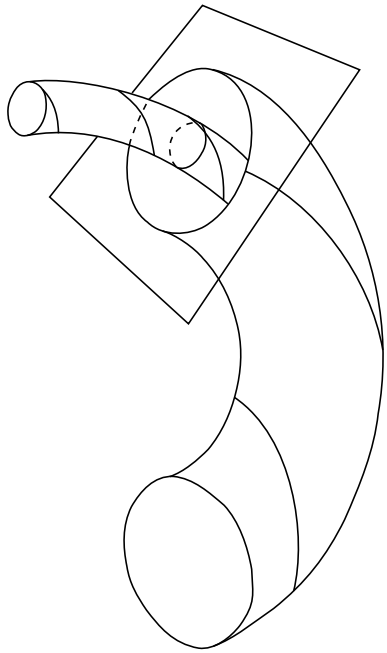


Figura 8.1: En el sistema integrable H_0 el espacio fase queda estructurado por toros formados por las trayectorias que constituyen una foliación.

pudiera ser que la estructura se destruyera, desapareciendo los toros y la foliación. Pudiera también ser que los toros únicamente se deformaran para valores suficientemente pequeños de ϵ, y que, al incrementar ϵ y con ello la deformación, terminaran rompiéndose. Veamos como elucidar esto.

Tomemos nuevamente el caso de dos grados de libertad, en el entendido que mucho de lo que para él obtengamos será extendible a sistemas con más grados de libertad. Para la hamiltoniana no perturbada H_0, que es integrable, las trayectorias estructuran el espacio fase como se ilustra en la figura ((8.1)), donde hemos, por razones de simplicidad en la ilustración, enderezado los toros y trazado la figura desde la perspectiva de su inmersión en un espacio de tres dimensiones, al que podemos considerar como una sección de energía constante E del espacio fase de cuatro dimensiones. Tomemos ahora, como se ilustra, un plano que intersecte la foliación, una trayectoria que comience en dicho plano y por tanto en uno de los toros, avanzará sobre ese toro y, después de haber completado un periodo correspondiente al ciclo perpendicular al plano, la trayectoria volverá al plano intersectándolo en algún punto; en general este punto será distinto del primero, pero

estará sobre el mismo círculo[1] que constituye la intersección del plano con el toro quecontiene la trayectoria. Como por cada punto del espacio fase pasa una y sólo una trayectoria y éstas deben dar vuelta sobre el toro que les corresponde, podemos, por este procedimiento, establecer un mapeo del plano en si mismo: a un punto cualquiera del plano le corresponde aquél en que la trayectoria que comienza en él intersecta al plano después de dar una vuelta; nótese que en la otra dirección del toro la trayectoria habrá dado vueltas también, pero no necesariamente un número entero de ellas; de hecho ese número no entero estará determinado por la proporción entre las dos frecuencias que corresponden al toro de que se trate y será un número racional si se trata de un toro conmensurable o de uno irracional si se trata de un toro inconmensurable; en el primer caso la trayectoria será periódica, en el segundo cuasiperiódica. A un mapeo construido de esta forma se le llama mapeo de Poincaré, denominación que se aplica en general a mapeos obtenidos por la intersección de las trayectorias de ecuaciones diferenciales con subespacios del espacio fase cuando esas trayectorias dan vueltas y por ello retornan a dicho subespacio para intersectarlo. Denotaremos al mapeo de Poincaré por $P_0 : D \to D$, donde hemos denotado con D la región del plano donde queda definido dicho mapeo y utilizamos el subíndice cero para recordar que se trata del mapeo generado por el sistema hamiltoniano H_0.

Este mapeo tiene una serie de propiedades derivadas de las propiedades de las trayectorias del sistema hamiltoniano. Es un mapeo uno a uno, diferenciable y que preserva área. Esto último es una consecuencia directa del hecho de que el flujo del sistema hamiltoniano preserva el volumen en el espacio fase.

Las iteraciones de este mapeo, esto es el resultado de aplicar el mapeo a un punto del plano, volverlo a aplicar al punto resultante de la primera aplicación y así sucesivamente, generan, para cada punto inicial en el plano, una serie de puntos que llamaremos la órbita que pasa por dicho punto inicial. Estas órbitas, que consisten de secuencias de puntos y no de curvas continuas, son las soluciones del sistema dinámico discreto generado por el mapeo de Poincaré, naturalmente sus propiedades reflejan las de las trayectorias del sistema hamiltoniano y por ello resulta útil estudiarlas.

En primer lugar notamos que como los toros intersectan al plano en círculos y las trayectorias permanecen sobre los toros, dado un punto del plano este pertenecerá también al círculo de intersección con alguno de los toros

[1] En realidad se trata de una curva cerrada y no necesariamente de un círculo, sin embargo podemos suponer que se ha aplicado un homeomorfismo que endereza los toros sin perdida de generalidad.

y, por tanto, al mapear el punto se obtendrá otro punto del mismo círculo y como esto ocurre para todas las iteraciones, la órbita de cualquier punto estará sobre el círculo correspondiente. Ahora si se trata de un toro conmensurable o resonante, la trayectoria será una curva que después de un cierto número de vueltas, digamos k, terminará cerrándose, pues se trata de un movimiento periódico; por tanto, después de k iteraciones la órbita generada por el mapeo de Poincaré regresará al punto inicial, esto es la órbita discreta será también periódica y consistirá de un número finito de puntos sobre el círculo, al seguir aplicando el mapeo se repetirá y repetirá la misma secuencia de puntos. Por otro lado, si se trata de un toro inconmensurable, la trayectoria en el espacio fase será cuasiperiódica y llenará densamente ese toro; en este caso la órbita discreta no se cerrará nunca, consistirá de una infinidad de puntos que llenan densamente al círculo correspondiente sin caer nunca en el mismo punto, a estas órbitas del mapeo las llamaremos cuasiperiódicas. Como las propiedades de periodicidad o cuasiperiodicidad de las trayectorias en el espacio fase dependen del toro en que se encuentran, las órbitas que comienzan en un mismo círculo sobre el plano tendrán también las mismas propiedades, por lo que podemos hablar de círculos periódicos y de círculos cuasiperiódicos. En cada caso podemos contemplar esto como un mapeo del círculo en si mismo generado por una rotación con ángulo fijo; este ángulo dependerá del toro de que se trate y será conmensurable con 2π si el toro es conmensurable y será inconmensurable con 2π si el toro es inconmensurable. En la figura (8.2) se han trazado sobre el plano un número grande de iteraciones del mapeo de Poincaré de un sistema integrable para dos condiciones iniciales que caen sobre toros conmensurables y dos que lo hacen sobre toros inconmensurables, los círculos correspondientes, esto es la intersección de los toros con el plano, quedan esbozados, en el caso periódico por un número finito de puntos que se irán repitiendo y repitiendo a medida que continuamos iterando y en el caso cuasiperiódico por un conjunto de puntos que nunca se repiten y que irán llenando densamente el círculo a medida que iteramos. Si trazáramos una infinidad de órbitas obtendríamos una infinidad de círculos concéntricos, algunos algunos estarían compuestos por órbitas periódicas y otros por cuasiperiódicas y, en general, los correspondientes a órbitas periódicas se intercalarían entre los cuasiperiódicos de la misma manera en que se intercalan los números racionales entre los irracionales, por tanto los círculos cuasiperiódicos ocuparían toda la medida y los periódicos una medida cero, aunque estos últimos formarían un subconjunto denso dentro del conjunto de todos los círculos.

Esta es la imagen que obtenemos cuando el sistema hamiltoniano es integrable, esto es cuando en la ecuación ((8.1)) ϵ es cero, ahora nos preguntamos

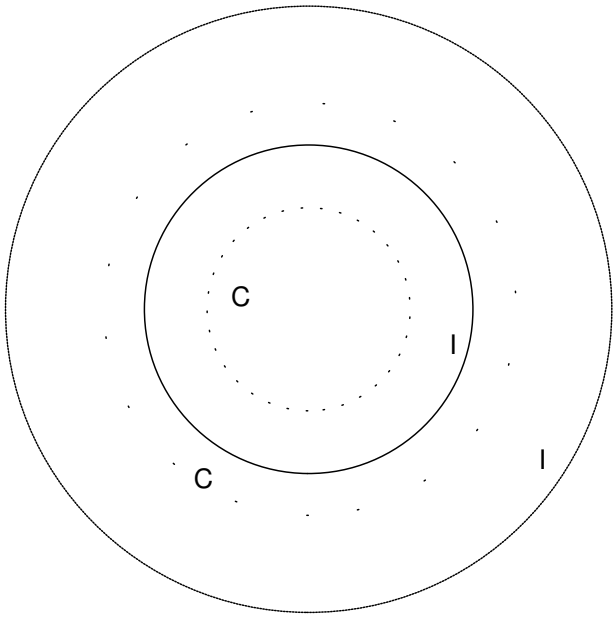


Figura 8.2: Cuatro órbitas del mapeo de Poincaré de un sitema integrable H_0. Las órbitas marcadas con una C son periódicas o conmensurables y están formadas por un número finito de puntos. Las órbitas marcadas con I son cuasiperiódicas o inconmensurables, están formadas por una infinidad de puntos, razón por la que, dadas las limitaciones en la resolución de estas gráficas, se ven continuas.

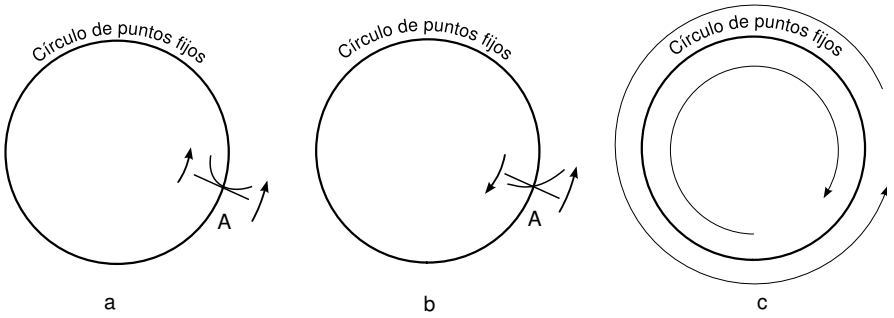


Figura 8.3: El mapeo $P_0^{(k)}$ en una vecindad de un toro resonante de período k.

que ocurre cuando ϵ deja de ser cero. Tomemos, en el caso integrable, uno de los toros conmensurables y su intersección con el plano que será un círculo también conmensurable. Digamos que la razón de las frecuencias sobre el toro es de la forma l/k donde l y k son enteros y primos relativos, esto es, las trayectorias sobre ese toro se cerrarán al dar k vueltas en una de las direcciones del toro, pues entonces se completarán exactamente l vueltas en la otra dirección y se cumple un período de la trayectoria. Desde el punto de vista del mapeo de Poincaré veremos una órbita discreta sobre el círculo consistente únicamente de k puntos, pues al mapear el punto inicial obtenemos sobre el círculo el punto resultante de hacer una rotación por un ángulo $2\pi l/k$, al ir iterando el punto de la órbita rotará cada vez por este mismo ángulo, por tanto al completarse k iteraciones regresamos al punto inicial y la órbita se cierra; todas las órbitas cuyo punto inicial esté sobre ese mismo círculo presentarán el mismo comportamiento.

Definamos ahora el mapeo $P_0^{(k)}$ como el mapeo que se obtiene al componer k veces el mapeo de Poincaré P_0, le llamaremos la potencia k del mapeo de Poincaré. Un punto sobre el círculo correspondiente al toro resonante l/k se mapea en aquel al que irá a parar después de iterar k veces el mapeo de Poincaré, o sea a él mismo, por tanto todos los punto de ese círculo son puntos fijos del mapeo $P_0^{(k)}$. Llegamos pues a la conclusión de que por cada toro conmensurable en un sistema integrable podemos construir, sobre la sección de Poincaré, un círculo completo constituido únicamente de puntos fijos de cierta potencia del mapeo de Poincaré.

En la figura (8.3) se muestra lo que ocurre en una vecindad de cualquier punto de ese círculo de puntos fijos de $P_0^{(k)}$. Tomemos un pequeño segmento de la recta radial que intersecta al círculo en un punto A, cuando aplicamos

el mapeo $P_0^{(k)}$ el círculo permanece invariante, pero el segmento no, excepto por el propio punto A. Dado que el mapeo es continuo podrían ocurrir dos situaciones como se muestra en la figura. En el caso (a) los puntos del segmento se han mapeado en una pequeña curva tangente al segmento en el punto A, esta situación por implicar una tangencia no será genérica, esto es ocurrirá únicamente en casos sumamente excepcionales. En oposición a lo anterior, en el caso (b) la pequeña curva es transversal al segmento original, esto será lo que ocurre en general. Vemos entonces que, en general, en una vecindad del círculo de puntos fijos, el mapeo $P_0^{(k)}$ rotará a los puntos que se encuentren en la parte de dentro del círculo en una dirección y a los que se encuentren afuera en la contraria. Un mapeo así se llama un mapeo de torcedura y la imagen genérica de este tipo de mapeos es la que se muestra en (c).

Veamos que ocurre cuando actúa la perturbación y la hamiltoniana es la señalada en ((8.1)); por continuidad, para ϵ suficientemente pequeña, las trayectorias en el espacio fase seguirán dando vueltas, por lo que será posible establecer un mapeo de Poincaré, llamémosle P_ϵ y construyamos a partir de él el mapeo $P_\epsilon^{(k)}$. Al aplicar el mapeo $P_\epsilon^{(k)}$ en una vecindad del círculo de puntos fijos de $P_0^{(k)}$ obtendremos lo que se muestra en la figura (8.4). Desde luego los puntos del círculo no serán puntos fijos de $P_\epsilon^{(k)}$, sin embargo, por continuidad, para ϵ suficientemente pequeña, $P_\epsilon^{(k)}$ seguirá siendo un mapeo de torcedura, por lo que, dada cualquier dirección radial, algún punto sobre ese radio cercano al círculo invariante de $P_0^{(k)}$ se mapeará en dirección radial al aplicarle $P_\epsilon^{(k)}$, esto es, no rotará ni en una, ni en la otra dirección. Tomemos ahora todas las direcciones, como en cada una de ellas hay un punto que se mapea únicamente de manera radial, todos ellos formarán una curva cerrada y continua que permanecerá cercana al antiguo círculo de puntos fijos de $P_0^{(k)}$ y que se mapeará radialmente en otra curva cerrada.

Pudiera ser que todos los puntos de esta curva cerrada, se mapearan alejándose del centro, o pudiera ser que ocurriera lo contrario que se mapearan acercándose, como se muestra en la figura (8.5); pero ambas situaciones son imposibles, ya que el mapeo de Poincaré P_ϵ proviene de un sistema hamiltoniano y por tanto preserva área; la única posibilidad es que algunos puntos se mapeen alejándose del centro y otros acercándose a él de tal suerte que el área interior de la curva original y la de la producida al aplicar el mapeo sean iguales, la única situación posible es la mostrada en la figura, necesariamente la curva formada por los puntos que se mapean radialmente se intersecta con la curva resultante al mapearla y, como son dos curvas cerradas, el número de puntos de intersección será par. Estos puntos de intersección no se mue-

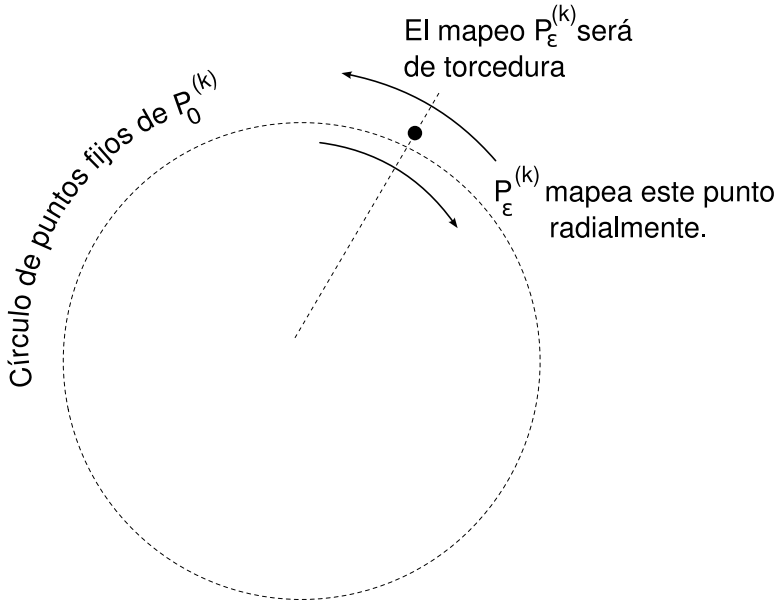


Figura 8.4: Dada una dirección existe un punto tal que al aplicarle un mapeo de torcedura se mapea radialmente.

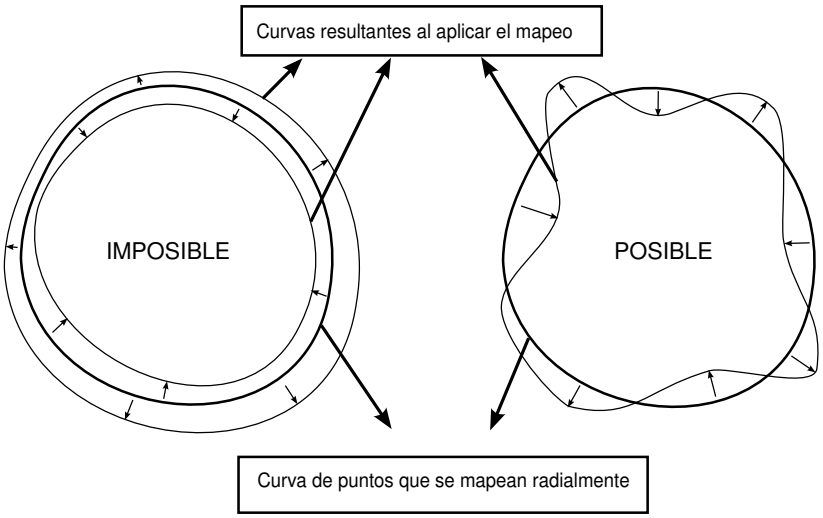


Figura 8.5: El mapeo de Poincaré preserva área, entonces mapea una curva cerrada en otra que la intersecta. Los puntos de intersección serán puntos fijos del mapeo.

ven ni radial ni angularmente al aplicar el mapeo, esto es, son puntos fijos de $P_\epsilon^{(k)}$ y por tanto pertenecen a órbitas de período k de P_ϵ, que, naturalmente, se corresponden con trayectorias periódicas en el espacio fase para la hamiltoniana perturbada ((8.1)). Este resultado que, desde luego, se aplica a cualquier mapeo de torcedura que preserva área, se conoce como el último teorema geométrico de Poincaré.

Regresando al punto de vista de la mecánica hemos encontrado algo de suma importancia. Al presentarse una perturbación por pequeña que sea de un sistema integrable acotado, la foliación por toros correspondiente se ve perturbada, por cada toro conmensurable aparece un número finito de trayectorias periódicas aisladas, o sea, en una vecindad de ellas no hay otras de ese período, esto implica que el toro conmensurable ya no existe, se rompió, y el flujo ya no tiene las propiedades que caracterizan a los sistemas integrables, no hay toros conmensurables y aparecen trayectorias periódicas aisladas, que no forman haces[2].

Pasemos a estudiar estas trayectorias periódicas aisladas que aparecen al presentarse una perturbación y romperse los toros conmensurables o resonantes. Desde el punto de vista del mapeo de Poincaré tendremos puntos fijos aislados del mapeo $P_\epsilon^{(k)}$; en una vecindad suficientemente pequeña de uno de estos puntos fijos el mapeo puede aproximarse linealmente y presentará el comportamiento general de un mapeo lineal. Así, si tomamos un punto como el indicado por A en la figura (8.6) tendremos que el punto fijo será estable en el sentido de que una órbita que comience cercana a él permanecerá cercana, ya que dará vueltas en su entorno; a este tipo de puntos fijos se les llama elípticos y, para la iteración de mapeos tienen características similares a los centros en las ecuaciones diferenciales. En contraste con esto, si el punto fijo es del tipo indicado por B en la figura, tendremos que al comenzar en una vecindad de él, en general, nos alejaremos siguiendo curvas hiperbólicas y el punto fijo será inestable, ya que las órbitas cercanas a él, en general no permanecerán cercanas, a este tipo de puntos fijos de un mapeo se les llama hiperbólicos. Podemos ver que los puntos fijos de $P_\epsilon^{(k)}$ serán de manera alternada elípticos e hiperbólicos. Veamos esto con detenimiento.

Pongamos un sistema cartesiano de coordenadas sobre el plano que constituye la sección de Poincaré, digamos (x, y), tal que su origen coincida con el punto fijo que estamos considerando. El mapeo $P_\epsilon^{(k)}$, en estas coordenadas,

[2]Desde luego esto puede verse como una demostración de la existencia de sistemas no integrables y además nos hace ver que, genéricamente hablando, dado un sitema integrable existe una vecindad de él en la que los sistemas no integrables son densos.

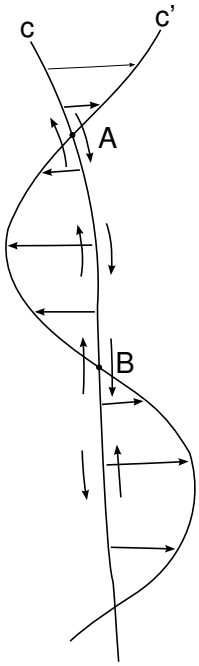


Figura 8.6: La curva c es la curva de puntos que se mapean radialmente por $P_\epsilon^{(k)}$ y la curva $c' = P_\epsilon^{(k)}(c)$. Las flechas indican la dirección en que se mueve un punto la aplicar el mapeo de torcedura $P_\epsilon^{(k)}$. El punto A es fijo y estable, el punto B es fijo pero inestable.

será

$$x_{i+1} = f(x_i, y_i), \tag{8.25}$$

$$y_{i+1} = g(x_i, y_i); \tag{8.26}$$

ya que el origen es el punto fijo tendremos que

$$f(0,0) = 0, \tag{8.27}$$

$$g(0,0) = 0, \tag{8.28}$$

por lo que en una vecindad del punto fijo podemos aproximar el mapeo por su linealización

$$\begin{pmatrix} x_{i+1} \\ y_{1+1} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \tag{8.29}$$

donde la matriz (a_{ij}) será

$$(a_{ij}) = \left.\frac{\partial(f,g)}{\partial(x,y)}\right|_0. \tag{8.30}$$

El comportamiento de las órbitas en un entorno del punto fijo dependerá de los eigenvalores de la matriz, en este caso, por tratarse de un sistema hamiltoniano que preserva área y orientación, el determinante de (a_{ij}) tiene que ser uno, por lo que, si los dos eigenvalores son reales deberán ser de la forma λ y $1/\lambda$ y, si son complejos conjugados, tendrán que estar sobre el círculo unitario y ser de la forma

$$e^{\pm i\lambda}. \tag{8.31}$$

En el primer caso, a no ser que λ sea cero o uno, casos que excluimos sin perdida de generalidad, un eigenvalor tendrá valor absoluto mayor que uno (digamos que es λ) y el otro lo tendrá menor que uno, por lo tanto al aplicar ((8.29)) a un punto sobre la recta generada por el eigenvector correspondiente a λ obtendremos un punto sobre la misma recta alejado por la razón λ y al aplicar el mapeo a un punto sobre la recta generada por el eigenvector correspondiente a $1/\lambda$ se acercará al punto fijo por una razón $1/\lambda$. Si se aplica el mapeo a un punto cualquiera, seguirá una órbita que simultáneamente se acerca en una dirección y se aleja en la otra por lo que caerá sobre una curva hiperbólica. Es interesante hacer ver la diferencia que hay entre los casos en que λ es positiva o negativa. En el primero la órbita se trazará siempre del mismo lado de las asíntotas y en el segundo brincará alternativamente de uno y otro lado.

En el caso en que los eigenvalores sean complejos conjugados el mapeo se trata de una rotación por un ángulo de λ radianes y las órbitas caerán sobre círculos, que se irán recorriendo de manera periódica o cuasiperiódica dependiendo de que λ sea conmensurable o inconmensurable con 2π.

Todo esto que hemos descrito para las órbitas cercanas a los puntos fijos de $P_\epsilon^{(k)}$ se traducirá de manera directa para el mapeo simple de Poincaré, en donde un punto fijo de $P_\epsilon^{(k)}$ es uno de los puntos que constituyen una órbita de período k, por lo que las órbitas sobre los círculos en el caso estable o las hipérbolas en el inestable se trazarán de manera sucesiva con las iteraciones en vecindades de los k puntos que constituyen la órbita periódica. Lo mismo podemos decir de las trayectorias correspondientes en el espacio fase, las órbitas periódicas del mapeo de Poincaré se corresponden con una trayectoria cerrada, esto es periódica, que da k vueltas en una dirección antes de cerrarse habiendo dado, en el mismo tiempo, l vueltas en la otra dirección. Las órbitas hiperbólicas y elípticas que se forman en una vecindad de los puntos fijos, se ven en el espacio fase como trayectorias que dan vueltas, pero que no se cierran; unas, las estables, permanecen sobre toros que dan k vueltas antes de cerrarse como una especie de manguera que se enrolla varias veces antes de cerrarse y las otras, las inestables, irán alejándose de la estrictamente periódica de manera hiperbólica. Naturalmente la formación de esos toros en torno a los puntos elípticos ocurre únicamente en una vecindad suficientemente pequeña del punto fijo donde predomina el mapeo linealizado.

En conclusión, una perturbación suficientemente pequeña de un sistema hamiltoniano integrable acotado conduce a la destrucción de los toros resonantes o conmensurables, las trayectorias periódicas que lo formaban se desorganizan y dejan de generar un toro, es más la mayoría de ellas dejan de ser periódicas con la excepción de un número par de ellas que conforman trayectorias periódicas aisladas del mismo período que las que formaban el toro destruido. La mitad de estas trayectorias serán estables y, vistas desde la perspectiva del mapeo de Poincaré, generan puntos periódicos elípticos, la otra mitad serán inestables y generan puntos periódicos hiperbólicos del mapeo. Todo esto se ilustra en la figura (8.7) para el caso en que se tenía período tres en el sistema integrable y únicamente se generan dos trayectorias periódicas de ese período, una estable y una inestable.

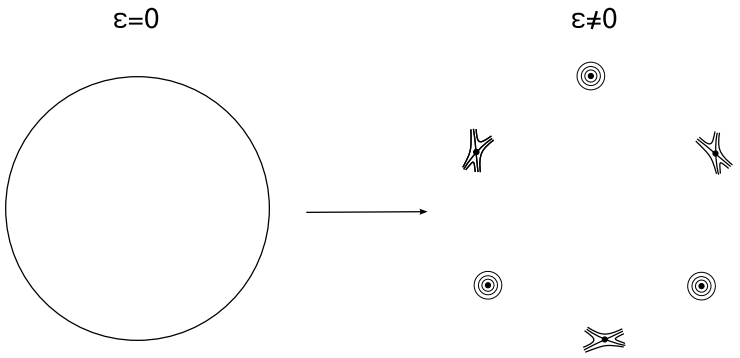


Figura 8.7: Un toro conmensurable de un sistema integrable se rompe en el momento en que se ve perturbado. En su lugar quedan trayectorias periódicas aisladas estables e inestables del mismo período que las que formaban el toro en el sistema integrable.

8.3. EL teorema de KAM.

En la sección anterior vimos que los toros conmensurables de un sistema hamiltoniano integrable se destruyen en el momento mismo en que se perturba el sistema y deja de ser integrable; pero no hemos dicho nada acerca de lo que ocurre con los toros inconmensurables, que ocupan la mayor parte del espacio fase; de hecho, en general, para el sistema integrable la medida del espacio fase ocupado por los toros inconmensurables es uno. En tiempos de Poincaré muchos pensaban que también se destruían en el instante mismo en que se perturbaba al sistema integrable, pero algunos pensaban que no. Encontrar solución a esta cuestión fue una de las principales motivaciones del trabajo de Poincaré, desafortunadamente no llegó a encontrar la respuesta a pesar de que este problema inspiró una buena parte de sus contribuciones a la mecánica y la matemática.

El problema del destino de los toros inconmensurables va más allá del contexto de la mecánica, en particular adquiere gran trascendencia desde la perspectiva de la mecánica estadística. Uno de los puntos de mayor interés en la física está en la conexión entre la teoría básica del movimiento, en nuestro caso la mecánica clásica, y el comportamiento de los sistemas macroscópicos compuestos por una multitud de partículas en movimiento, cada una de ellas respondiendo a la teoría básica. En especial es de destacar el hecho de que la experiencia señala que los procesos macroscópicos son irreversibles, mientras que las leyes básicas en el mundo microscópico son totalmente reversibles. ¿Cómo se genera la irreversibilidad macroscópica cuando la teoría básica es

reversible?.

Desde la perspectiva macroscópica el estado de un sistema está caracterizado por unas cuantas variables de estado, básicamente las manejadas por la termodinámica, que nos ofrecen una caracterización física del sistema. Por otro lado, el mismo sistema visto desde la perspectiva microscópica requiere de un sinnúmero de variables, del orden de 10^{25}, por lo que no hay duda que a un estado macroscópico le corresponden muchísimos estados microscópicos. Los estados macroscópicos presentan una tendencia temporal que irreversiblemente los lleva a un estado de equilibrio macroscópico, en éste todas las variables macroscópicas permanecen homogéneas y constantes y la entropía se hace máxima; nunca se ha observado un sistema macroscópico libre y cerrado que se aparte de esto, por ejemplo uno en el que la entropía disminuya alejándose del valor máximo compatible con el propio sistema. Por otro lado, desde la perspectiva microscópica el sistema nunca está estático, las más de 10^{25} variables cambian incesantemente con el tiempo y este cambio es reversible, basta con invertir en un instante dado los signos de las velocidades generalizadas o de los impulsos, para que el sistema recorra su trayectoria en dirección contraria; de donde sale entonces la irreversibilidad termodinámica.

A partir de las brillantes ideas introducidas por Boltzman se induce una respuesta; en forma esquemática diremos que el número de microestados que corresponden a un estado de equilibrio macroscópico es inmensamente grande en comparación con los que corresponden a situaciones fuera de equilibrio, esto es una simple consecuencia del inmenso número de grados de libertad del microsistema; también tenemos el hecho de que cuando preparamos un sistema para que quede fuera de equilibrio lo hacemos siempre manipulando las variables macroscópicas y no la inmensidad de variables microscópicas, por lo que no tenemos una cualquiera de las condiciones iniciales microscópicas congruentes con la preparación del sistema, resultando favorecidas las que al paso del tiempo llevan a estados de equilibrio macroscópico. Claro está que esto no es una deducción formal y cuestiones como el teorema de recurrencia hacen necesaria la introducción de ciertos postulados adicionales o ciertas restricciones en toda teoría estadística de los fenómenos macroscópicos; argumentos como los enunciados únicamente hacen plausibles estos postulados.

Ahora bien si el sistema, desde la perspectiva microscópica fuese integrable, toda trayectoria quedaría confinada a un toro de dimensión igual al número de grados de libertad y sería periódica o cuasiperiódica, por lo que en general se induciría una variación periódica o cuasiperiódica de las variables macroscópicas y el sistema sería reversible macroscópicamente ha-

blando. Es más, podemos ver que muchos de estos toros están compuestos de estados microscópicos que desde la perspectiva macroscópica no son estados de equilibrio, por lo que no habría tendencia al equilibrio. Esto no es grave, puesto que con seguridad un sistema físico real de 10^{25} grados de libertad no es integrable, pero si ocurriera que en los sistemas no integrables, de alguna manera, las trayectorias permanecieran sobre variedades de dimensión considerablemente menor que la del espacio fase, ocurriría algo similar y tendríamos dificultades para explicar la tendencia al equilibrio macroscópico. De aquí la importancia de resolver la pregunta sobre las trayectorias en los sistemas no integrables y, para comenzar, sobre el destino de las trayectorias que en un sistema integrable conformaban los toros inconmensurables al dejar dicho sistema de ser integrable vía una perturbación.

En general se pensaba en tiempos de Poincaré que la foliación por toros se rompía totalmente, esto es, que les ocurre a los toros cuasiperiódicos algo similar a lo que les ocurre a los conmensurables y desaparece totalmente la organización en variedades de dimensión pequeña formadas por haces de trayectorias; esto implicaría que las trayectorias en general explorarían una parte considerable de la variedad de energía constante, esto es, serían en ella ergódicas. Como una inmensa mayoría de los microestados corresponden a estados de equilibrio, la ergodicidad implica que el sistema estaría en estados de equilibrio durante tiempos inmensamente mayores que en estados fuera de equilibrio. Así las ideas de Boltzmann inducirían un entendimiento de la irreversibilidad macroscópica.

Naturalmente existía la opinión contraria, esto es que al menos para sistemas cercanos a uno integrable la organización de las trayectorias en el espacio fase impide la ergodicidad sobre la variedad de energía constante. La ausencia de una demostración en una dirección u otra o de ejemplos solubles o numéricamente atacables, simplemente dejaba el asunto a nivel de opinión.

Fermi, entre muchos otros, se planteó este problema en los años veinte de siglo XX, pensó que los toros cuasiperiódicos se rompían y el sistema se volvía ergódico en cuanto dejaba de ser integrable, incluso creyó tener una demostración de ello.

A mediados de los años 50 el propio Fermi, en colaboración con Pasta y Ulam, aprovechando el acceso que tenían a la computadora ENIAC de Los Álamos, atacaron esta cuestión haciendo por primera vez uso heurístico de una computadora para atacar un problema de la física básica. Cuando un sistema está en equilibrio macroscópico su energía total se distribuye uniformemente entre sus grados de libertad, lo que se conoce como el teorema de equipartición de la energía; si tenemos un sistema fuera de equilibrio en el que la energía se encuentra concentrada en algunos de sus grados de libertad,

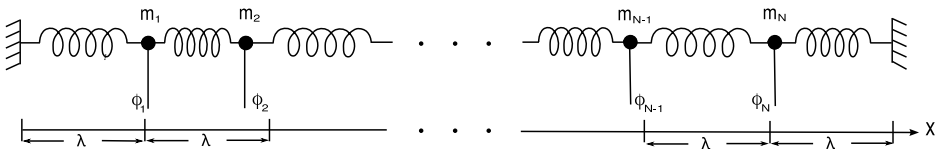


Figura 8.8: El sitema de masas y resortes de Fermi, Pasta y Ulam.

al tender éste con el tiempo al equilibrio, veremos que la energía se va difundiendo hasta llegar a una distribución homogénea en los grados de libertad; la ocurrencia de esto nos mostraría la tendencia al equilibrio. Fermi, Pasta y Ulam se propusieron ver esto para un sistema de muchas partículas pero suficientemente sencillo como para que ENIAC pudiera atacarlo numéricamente. Tomaron un sistema de partículas en una dimensión entre las que se intercalan resortes, esto es el sistema que estudiamos en el capítulo cuatro. Tomaron en una de las simulaciones 32 partículas y en otra 64, todas ellas de la misma masa, y los resortes también todos iguales, amarraron con resortes las partículas de los extremos a paredes rígidas, lo que fija las condiciones a la frontera. El sistema se ve entonces como se muestra en la figura (8.8).

La lagrangiana está dada en la ec*** y de ella se desprende la hamiltoniana

$$H_0 = \sum_j \frac{p_j^2}{2m_j} + \frac{1}{2} h_j \sum_j (\phi_{j+1} - \phi_j)^2, \tag{8.32}$$

la que, al hacer todas las masas y constantes de resorte iguales a uno, se reduce a

$$H_0 = \sum_j \frac{1}{2} \left[p_j^2 + (\phi_{j+1} - \phi_j)^2 \right]. \tag{8.33}$$

Este sistema es integrable, pues es lineal, y además lo podemos estudiar en coordenadas normales. Hay un modo normal por cada grado de libertad, por lo que podemos decir que cuando el sistema está en equilibrio la energía total estará repartida uniformemente entre los modos normales. Naturalmente si comenzamos con una distribución de energía no homogénea, digamos que ponemos toda la energía en un solo modo normal, el sistema permanecerá con esa misma distribución, los modos no están acoplados por lo que no hay transferencias de energía de un modo a otro. Sin embargo si introducimos una perturbación y el sistema se hace no integrable, aparecerá un acoplamiento entre los modos[3] y al transcurrir el tiempo la energía se irá transfiriendo de

[3]Estrictamente ablando ya no habría modos normales, el sistema ya no es lineal debido a la perturbación, sin embargo podemos continuar utilizando las coordenadas normales y

un modo a otro; si hubiera tendencia al equilibrio, después de un tiempo, tendríamos una distribución estática y homogénea de la energía entre los modos. Naturalmente esto es de esperarse si al hacerse no integrable el sistema desapareciera toda la estructura en foliaciones que se tenía en el caso integrable, como pensaba Fermi. Podemos, en principio, utilizando este modelo ver la destrucción de la foliación y la tendencia al equilibrio; de hecho FPU esperaban ver como tiende al equilibrio el sistema analizando la evolución temporal de la distribución de la energía entre los grados de libertad. Para hacer esto simplemente introducen una pequeña perturbación por medio de adicionar un término no lineal al potencial de cada resorte. Veamos esto con más detenimiento.

Los modos normales del sistema no perturbado serán aquellos movimientos en los que todas las partículas oscilen con la misma frecuencia, las ecuaciones de movimiento son

$$\ddot{\phi}_l = \phi_{l-1} - 2\phi_l + \phi_{l+1} \tag{8.34}$$

donde, siguiendo la numeración que se ilustra en la figura, tendremos como condiciones a la frontera

$$\phi_0 = \phi_{N+1} = \dot{\phi}_0 = \dot{\phi}_{N+1} = 0; \tag{8.35}$$

en uno de lo modos normales las soluciones deben ser de la forma

$$\phi_l(t) = a_l \text{sen}\omega t, \tag{8.36}$$

sustituyendo encontramos que

$$-a_l\omega^2 = a_{l-1} - 2a_l + a_{l+1}. \tag{8.37}$$

Proponemos una solución de estas ecuaciones algebraicas de la forma

$$a_l = C\text{sen}Bl\lambda, \tag{8.38}$$

lo que al substituir y hacer algunas manipulaciones trigonométricas nos lleva a que

$$\omega = 2\text{sen}^2\frac{B\lambda}{2}, \tag{8.39}$$

pero las condiciones a la frontera nos obligan a poner que $a_0 = a_{N+1} = 0$ por lo que

$$0 = C\text{sen}B\lambda(N+1), \tag{8.40}$$

---

calcular la energía en cada una de ellas, esto será una forma de ver la energía por grado de libertad.

o sea

$$B\lambda(N+1) = k\pi \tag{8.41}$$

y

$$a_l = C\text{sen}\frac{lk\pi}{N+1}, \tag{8.42}$$

por lo que

$$\omega_k^2 = 4\text{sen}^2\frac{k\pi}{2(N+1)} \tag{8.43}$$

y

$$\phi_l^{(k)} = C\text{sen}\frac{kl\pi}{N+1}\text{sen}\omega_k t \qquad k = 1, \ldots, N, \tag{8.44}$$

donde con $\phi_l^{(k)}$ hemos denotado a la coordenada ϕ_l cuando el sistema está en el modo normal k y C es una constante arbitraria que usaremos para normalizar.

En coordenadas normales

$$Q^j = \sum_l \alpha_l^j \phi_l \tag{8.45}$$

el modo normal k se da cuando $Q^k = A_k\text{sen}\omega_k t$ y $Q^{j\neq k} = 0$, por lo que, utilizando el hecho de que

$$\sum_{j=1}^{N} \text{sen}\frac{kj\pi}{N+1}\text{sen}\frac{jl\pi}{N+1} = \frac{N+1}{2}\delta_{kl}, \tag{8.46}$$

llegamos a unas coordenadas normales de la forma

$$Q^k = \frac{2}{N+1}\sum_l \text{sen}\frac{lk\pi}{N+1}\phi_l, \tag{8.47}$$

en ellas las ecuaciones de movimiento serán

$$\ddot{Q}^k + \omega_k^2 Q^k = 0, \tag{8.48}$$

que nos hacen ver, como siempre en modos normales, que los N grados de libertad están desacoplados y que por tanto el movimiento será cuasiperiódico o periódico dependiendo de la inconmensurabilidad o conmensurabilidad de las frecuencias ω_k, lo que corresponde a un sistema integrable. Naturalmente en este sistema, como en cualquiera que fuese integrable, las trayectorias permanecen siempre sobre toros de N dimensiones y no presentará una tendencia al equilibrio, pues no es, ni con mucho, ergódico sobre la superficie

de energía constante. Así, si como ejemplo, utilizamos condiciones iniciales que corresponden a excitar un único modo normal, la energía no se difundirá entre los otros modos normales, ya que no hay acoplamiento entre ellos y no se llegará nunca a una equipartición de la energía, de hecho cualquier distribución inicial de la energía entre los modos normales permanecerá invariante ante el movimiento del sistema.

Por otro lado si introducimos una perturbación en el sistema que lo haga no integrable, entonces aparecerá un acoplamiento entre los modos normales, o lo que es lo mismo entre los grados de libertad, y si se rompe totalmente la estructura del espacio fase en foliaciones podemos esperar un proceso de intercambio de energía entre los diversos grados de libertad, en este caso vistos como las coordenadas normales del sistema integrable, y una tendencia irreversible a la equipartición de la energía entre ellos. Esto es lo que Fermi, Pasta y Ulam (FPU) esperaban ver, es más Fermi pensaba que había demostrado que esto ocurría en general.

Para introducir una perturbación FPU adicionan al potencial cuadrático de los resortes una corrección cúbica o una cuártica, decimos corrección porque FPU pensaban en simular un cristal unidimensional y serían las primeras correcciones a la componente armónica del potencial entre los átomos del cristal. Así la hamiltoniana ((8.33)) se substituye por

$$H = \sum_j \frac{1}{2}\left[p_j^2 + (\phi_{j+1} - \phi_j)^2 + \epsilon_1(\phi_{j+1} - \phi_j)^3 + \epsilon_2(\phi_{j+1} - \phi_j)^4\right], \quad (8.49)$$

donde ϵ_1 y ϵ_2 son parámetros pequeños que nos miden la intensidad de la perturbación. En las coordenadas normales del sistema integrable (Q^k, P_k) esta hamiltoniana se expresa como

$$H = \sum_k \frac{1}{2}\left[P_k^2 + \omega_k^2 Q^{k2}\right] + \epsilon_1 \sum_{ijk} A_{ijk} Q^i Q^j Q^k + \epsilon_2 \sum_{ijkl} B_{ijkl} Q^i Q^j Q^k Q^l, \quad (8.50)$$

donde los coeficientes A_{ijk} y B_{ijkl} se calculan a partir de la transformación de coordenadas ((8.47)). Esta expresión nos muestra a la hamiltoniana como una hamiltoniana integrable perturbada y la aparición de un acoplamiento pequeño con ϵ_1 o ϵ_2 entre los grados de libertad.

En realidad FPU hacen dos cálculos uno para el caso cúbico y otro para el caso cuártico, los resultados para ambos son similares. Se pone como condición inicial toda la energía en un sólo modo normal, el primero, esto es $Q^k = P_k = 0$, excepto para $k = 1$ y se deja correr en la computadora la solución de las ecuaciones diferenciales del sistema. A medida que transcurre el

tiempo y se va avanzando en la solución se va calculando la energía contenida en cada coordenada normal, de esta manera podemos ver como evoluciona la distribución energía entre los grados de libertad; recordamos que lo que esperaban era una evolución hacia la equipartición a medida que el sistema tiende al equilibrio. Los resultados se presentan en la figura (8.9) adaptada del propio trabajo de FPU. En ella se gráfica la energía contenida en los cinco primeros modos normales en función del tiempo transcurrido. Podemos apreciar que al ir pasando el tiempo, como se esperaba, la energía se difunde como lo muestran las gráficas, pero vemos un fenómeno inesperado pues a partir del paso 25000 la energía en vez de continuar difundiéndose regresa a concentrarse y en el tiempo correspondiente al paso 29000, en gran parte se ha reconcentrado en el primer modo, es decir retorna en gran medida a la condición inicial. Si se continua el cálculo se ve que este fenómeno se repite, es más después de haber transcurrido varios retornos la concentración en el primer modo es aun más pronunciada que en el primer retorno, esto se repite y se repite y ocurre tanto para el caso cúbico como para el cuártico.

El comportamiento del sistema es muy distinto al que esperaban FPU, no hay tendencia a la equipartición de la energía, más bien se observa lo que pudiera ser un fenómeno cuasiperiódico, esto a pesar de que el sistema se ha hecho no integrable y de tener muchos grados de libertad; desde luego[4] no se observa una tendencia al equilibrio. Ulam relata como Fermi se sorprendió de estos resultados y los calificó como de un verdadero descubrimiento de importancia científica, aunque desafortunadamente no llegó a desarrollarlos e interpretarlos, pues murió unas semanas antes de que estos calculos estuvieran totalmente terminados. En adición a esto, esta última contribución de Fermi se mantuvo oculta durante una década, pues, como los cálculos habían sido hechos en la computadora ENIAC de Los Álamos, el artículo en que se pretendían publicar fue 'clasificado' y permaneció archivado en la bóveda de los grandes secretos.

En una perspectiva distinta desde el siglo XIX se propusieron diversas formas para estudiar los sistemas no integrables; entre ellos, la teoría de perturbaciones tiene un papel destacado. En general, por una teoría de perturbaciones entendemos un método que nos permita estudiar las propiedades de las soluciones de un sistema de la forma

$$H = H_0 + \epsilon H_1 + \frac{\epsilon^2}{2} H_2 + \cdots , \tag{8.51}$$

donde H_0 es integrable y ϵ un parámetro pequeño, a partir de un desarrollo de

[4]FPU lo hicieron para 32 y para 64 partículas, posteriormente se han hecho cálculos para muchas más partículas y diversas formas de la interacción con resultados similares.

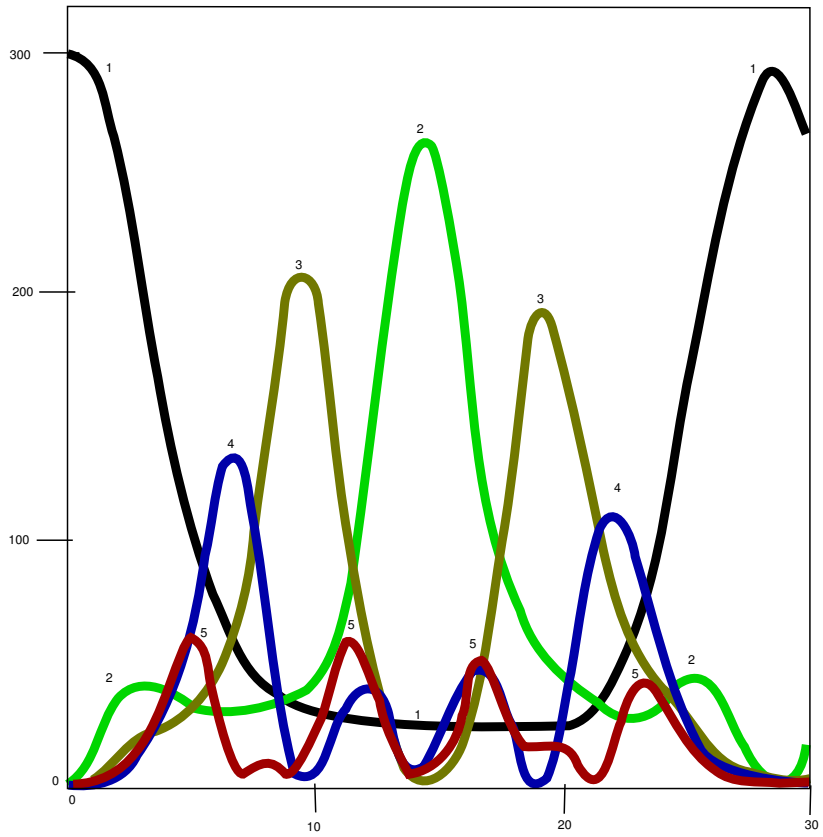


Figura 8.9: En las gráficas se muestran los resultados de Fermi, Pasta y Ulam para el caso de 32 osciladores con una perturbación cúbica. En las abscisas tenemos el número de pasos que ha realizado la computadora en miles, equivalente al tiempo trascurrido; en las ordenadas tenemos la energía total en unidades arbitrarias; las gráficas muestran la energía contenida en los cinco primeros 'modos normales'. Vemos que inicialmente toda la energía estaba en el modo 1, conforme va pasando el tiempo la energía se va compartiendo entre los modos, pero podemos ver claramente como aproximadamente en el paso 29000 se ha concentrado nuevamente en el modo 1. Esta gráfica se ha adaptado directamente de la original obtenida por FPU.

la solución buscada en series de potencias de ϵ. Así se propone una solución del sistema ((8.51)) de la forma

$$q(t) = q_0(t) + \epsilon q_1(t) + \frac{\epsilon^2}{2} q_2(t) + \cdots , \tag{8.52}$$

donde $q_0(t)$ es solución del sistema integrable H_0. Diversas formas de la teoría de perturbaciones presentan diversos métodos para ir encontrando las sucesivas aproximaciones $q_1, q_2, \ldots$, en el entendido que, al ser ϵ suficientemente pequeña, podemos cortar el desarrollo en un cierto nivel y obtener la solución con el grado de aproximación que requiramos. Cuando nos quedamos hasta el término q_1 decimos que hemos considerado perturbaciones a primer orden, si incluimos el término q_2 hablamos de segundo orden y así sucesivamente, claro esta, esto tiene sentido únicamente si la serie que obtengamos es convergente, de lo contrario no podemos usar la idea de las perturbaciones o, a lo más, ésta podría ser de cierta utilidad para estudiar la solución a tiempos suficientemente cortos como para que la no convergencia no distorsione significativamente los resultados, pero no para concluir cuestiones generales sobre el sistema.

La teoría de perturbaciones comienza a desarrollarse ya en el siglo XIX para encontrar soluciones aproximadas de sistema de importancia; por ejemplo para predecir la órbita de un planeta cuando se toma en consideración la influencia de otros planetas; de hecho se utilizaron cálculos de esta naturaleza para el descubrimiento de Urano y muy destacadamente de Neptuno y Plutón. En la actualidad se utilizan para asuntos como el cálculo de las órbitas de las naves espaciales o para el estudio del comportamiento del sistema de tres cuerpos o, claro está, en el asunto que estamos desarrollando, esto es el comportamiento general de los sistemas no integrables cercanos a uno integrable.

La teoría de perturbaciones forma un apartado sumamente amplio y especializado de la mecánica, aquí únicamente trataremos algunos aspectos que nos servirán para ilustrar las dificultades a que nos lleva y para permitirnos apreciar los resultados que sobre el tema de nuestro interés se han logrado.

De entre las muchas formas en que se ataca la teoría de perturbaciones exploraremos únicamente dos: la teoría secular y la teoría canónica de perturbaciones.

La teoría secular, parte directamente de las ecuaciones de movimiento para construir la solución en serie. Tomemos como ejemplo ilustrativo el caso de un oscilador armónico perturbado. El sistema no perturbado es el

oscilador armónico, esto es

$$H_0 = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 \tag{8.53}$$

y como término de perturbación consideramos un potencial cuártico

$$H_1 = \frac{1}{4}Cx^4, \tag{8.54}$$

por tanto el sistema perturbado será

$$H = H_0 + \epsilon H_1 = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 + \epsilon\frac{1}{4}Cx^4. \tag{8.55}$$

Por ser este un sistema en un grado de libertad es reducible a cuadraturas, pero como lo estamos utilizando como ilustración, lo atacaremos utilizando perturbaciones. Para ello substituimos un desarrollo en serie directamente en las ecuaciones de movimiento.

La ecuación de movimiento del sistema no perturbado es

$$\ddot{x} = -\omega_0^2 x \tag{8.56}$$

con

$$\omega_0 = \sqrt{\frac{k}{m}}, \tag{8.57}$$

y la ecuación de movimiento del sistema completo será

$$\ddot{x} + \omega_0^2 x + \epsilon\frac{C}{m}x^3 = 0. \tag{8.58}$$

Llamando $x_0, x_1, x_2, \ldots$ a las aproximaciones de orden cero, uno, dos, etc. expresaremos la solución de sistema perturbado por

$$x = x_0 + \epsilon x_1 + \frac{\epsilon^2}{2}x_2 + \cdots, \tag{8.59}$$

substituyendo ((8.59)) en ((8.58)) obtenemos

$$\begin{aligned}\ddot{x}_0 + \epsilon\ddot{x}_1 + \frac{\epsilon^2}{2}\ddot{x}_2 + \cdots &= \omega_0^2\left(x_0 + \epsilon x_1 + \frac{\epsilon^2}{2}x_2 + \cdots\right) \\ &\quad + \frac{C}{m}\left(\epsilon x_0^3 + 3\epsilon^2 x_0^2 x_1 + \cdots\right),\end{aligned}$$

igualando los términos que presentan la misma potencia de ϵ tendremos

$$\ddot{x}_0 + \omega_0^2 x_0 = 0, \tag{8.60}$$

$$\ddot{x}_1 + \omega_0^2 x_1 + \frac{C}{m} x_0^3 = 0, \tag{8.61}$$

$$\ddot{x}_2 + \omega_0^2 x_2 + 6\frac{C}{m} x_0^2 x_1 = 0, \tag{8.62}$$

$$\vdots \tag{8.63}$$

La primera de las ecuaciones de esta jerarquía es la del oscilador armónico sin perturbar y su solución es

$$x_0 = A\cos(\omega_0 t + \beta). \tag{8.64}$$

Utilizando esta solución en la segunda de las ecuaciones de la jerarquía obtenemos una ecuación diferencial no autónoma para x_1

$$\ddot{x}_1 + \omega_0^2 x_1 + \frac{C}{m} A^3 \cos^3(\omega_0 t + \beta) = 0, \tag{8.65}$$

una de cuyas soluciones es[5]

$$x_1(t) = \frac{CA^3}{8m\omega_0^2}\left\{3\omega_0 t \sin(\omega_0 t + \beta) + \frac{1}{4}\left[\cos(\omega_0 t + \beta) - \cos(3\omega_0 t + 3\beta)\right]\right\} \tag{8.66}$$

que nos proporciona la aproximación a primer orden a la solución de ((8.58)). Ahora podemos substituir x_0 y x_1 en la tercera ecuación de la jerarquía y obtener una ecuación para x_2 cuya solución nos proporciona el segundo orden de aproximación. Siguiendo este procedimiento podemos encontrar el orden de aproximación que queramos. Aparentemente hemos resuelto, con el grado de aproximación que se desee, el sistema perturbado. Sin embargo se presenta una dificultad muy seria pues, si nos fijamos en el primer término de la derecha de ((8.66)) vemos que es una oscilación cuya amplitud crece indefinidamente con el tiempo debido a la presencia del término $3\omega_0 t \sin(\omega_0 t + \beta)$. Esto nos hace ver que la serie que planteamos no converge y por tanto resulta inútil.

Este tipo de términos se presenta con frecuencia en los métodos perturbativos y reciben el nombre de términos seculares; constituyen una de las dificultades más significativas de la teoría de perturbaciones, dificultad que resulta muchas veces insuperable.

[5] Hemos dejado que A y B sean las constantes con las que se establecen las condiciones iniciales.

En el ejemplo que estamos analizando, por su simplicidad y por ser integrable, se puede encontrar un procedimiento para evitar los términos seculares, veamos esto. La ecuación ((8.65)) es equivalente a la de un oscilador armónico de frecuencia ω_0 forzado por un término de la forma

$$-\frac{c}{m}A^3\cos^3(\omega_0 t+\beta), \tag{8.67}$$

pero usando la igualdad trigonométrica

$$\cos^3\alpha=\frac{1}{4}(3\cos\alpha+\cos 3\alpha), \tag{8.68}$$

vemos que consiste en un oscilador armónico forzado periódicamente con un término de su propia frecuencia, por lo que entra en resonancia; esto explica el término divergente. Si pudiéramos evitar esta resonancia podríamos encontrar una serie convergente. La resonancia se debe a que utilizamos la frecuencia del sistema no perturbado como base para calcular la serie perturbativa, pero sabemos que la frecuencia de la oscilación real correspondiente al sistema perturbado es distinta de ω_0, de hecho es dependiente de la amplitud; Poincaré procedió a tomar en consideración esto introduciendo una serie para la frecuencia del sistema perturbado ω:

$$\omega^2=\omega_0^2+\omega_1\epsilon+\frac{1}{2}\omega_2\epsilon^2+\cdots, \tag{8.69}$$

donde las cantidades ω_i deben determinarse del propio proceso perturbativo; veamos esto. La ecuación de movimiento (8.58) la podemos escribir como

$$\ddot{x}+(\omega^2-\epsilon\omega_1-\frac{1}{2}\epsilon^2\omega_2-\cdots)x+\epsilon\frac{C}{m}x^3=0, \tag{8.70}$$

al substituir el desarrollo (8.59) y agrupar términos en las potencias de ϵ obtenemos la jerarquía de ecuaciones

$$\ddot{x}_0+\omega^2x_0=0 \tag{8.71}$$

$$\ddot{x}_1+\omega^2x_1-\omega_1x_0+\frac{C}{m}x_0^3=0 \tag{8.72}$$

$$\vdots \tag{8.73}$$

La solución de la primera ecuación es

$$x_0=A\cos(\omega t+\beta), \tag{8.74}$$

y al substituir en la segunda ecuación de la jerarquía obtenemos una ecuación para x_1

$$\ddot{x}_1 + \omega^2 x_1 + \left(\frac{3CA^3}{4m} - \omega_1 A\right)\cos(\omega t + \beta) + \frac{CA^3}{4m}\cos(3\omega t + 3\beta) = 0, \quad (8.75)$$

donde ya se utilizó la igualdad (8.68). Vemos que tenemos nuevamente una resonancia, pero ahora tenemos a nuestra disposición la cantidad ω_1 para eliminarla, esto es hacemos

$$\omega_1 = \frac{3CA^2}{4m}, \quad (8.76)$$

con lo que se elimina la resonancia, así podemos obtener una expresión para la primera aproximación x_1 sin términos seculares, adicionalmente obtenemos una primera aproximación a la frecuencia del sistema perturbado puesto que de (8.69) encontramos que

$$\omega \cong \left[\omega_0^2 + \frac{3CA^2}{4m}\epsilon\right]^{\frac{1}{2}}, \quad (8.77)$$

que nos hace ver como la frecuencia depende de la amplitud del movimiento. Continuando con un proceso similar podríamos evitar las resonancias en los ordenes superiores, logrando con ello el resultado deseado.

En general para sistemas en más de un grado de libertad aparecen términos seculares que no son tan fácilmente evitables, esto debido básicamente a la presencia de términos en resonancia entre los distintos grados de libertad, por ello la teoría secular de perturbaciones resulta poco útil para llegar a conclusiones teóricas generales sobre los sistemas no integrables cercanos a uno integrable.

Otra forma de la teoría de perturbaciones hace un uso intenso de la estructura canónica o hamiltoniana de la mecánica, es la llamada teoría canónica de perturbaciones. En ella partimos del hecho de que un sistema integrable y acotado puede estudiarse en variables de acción y ángulo y al hacerlo la hamiltoniana queda en función únicamente de las variables de acción.

Tomemos un sistema

$$H = H_0 + \epsilon H_1 + \epsilon^2 H_2 + \cdots, \quad (8.78)$$

donde H_0 es integrable y que presenta una región acotada en el espacio fase. Las variables de acción y ángulo para H_0, (J, θ), conforman un sistema de

coordenadas canónicas en el que es totalmente valido expresar la hamiltoniana perturbada, tendremos

$$H(J,\theta) = H_0(J) + \epsilon H_1(J,\theta) + \epsilon^2 H_2(J,\theta) + \cdots , \tag{8.79}$$

donde $H_1, H_2, \ldots$ son funciones periódicas de período 2π en todos los ángulos θ^i, esto debido al carácter angular de las θ y a que la hamiltoniana debe ser una función continua y univaluada.

Planteemos este sistema en términos de la ecuación de Hamilton Jacobi. En este caso, al no haber una dependencia explicita del tiempo, podemos usar la ecuación independiente del tiempo (eq ***)

$$H_0\left(\frac{\partial W}{\partial\theta^1},\ldots,\frac{\partial W}{\partial\theta^n}\right)+\epsilon H_1\left(\frac{\partial W}{\partial\theta^1},\ldots,\frac{\partial W}{\partial\theta^n},\theta^1,\ldots,\theta^n\right)+\cdots = E_0+\epsilon E_1+\cdots , \tag{8.80}$$

donde hemos expresado la constante E en términos de su desarrollo en ϵ.

Ahora usamos la idea básica de la teoría de perturbaciones, planteamos una solución de esta ecuación como una serie del parámetro pequeño ϵ, esto es pongamos

$$W = W_0 + \epsilon W_1 + \cdots , \tag{8.81}$$

y busquemos la solución orden por orden en ϵ. Substituyendo ((8.81)) en ((8.80)) obtenemos

$$H_0\left(\frac{\partial(W_0+\epsilon W_1+\cdots)}{\partial\theta^1},\ldots,\frac{\partial(W_0+\epsilon W_1+\cdots)}{\partial\theta^n}\right)+\cdots \tag{8.82}$$

$$+\epsilon H_1\left(\frac{\partial(W_0+\epsilon W_1+\cdots)}{\partial\theta^1},\ldots,\frac{\partial(W_0+\epsilon W_1+\cdots)}{\partial\theta^n},\theta^1,\ldots,\theta^n\right) \tag{8.83}$$

$$= E_0 + \epsilon E_1 + \cdots \tag{8.84}$$

desarrollando en serie tendremos

$$H_0\left(\frac{\partial W_0}{\partial\theta^1},\ldots,\frac{\partial W_0}{\partial\theta^n}\right)+\epsilon\left[\sum_j \frac{\partial H_0}{\partial J_j}\frac{\partial W_1}{\partial\theta^j} + H_1\left(\frac{\partial W_0}{\partial\theta^1},\ldots,\frac{\partial W_0}{\partial\theta^n},\theta^1,\ldots,\theta^n\right)\right]+\cdots \tag{8.85}$$

$$= E_0 + \epsilon E_1 + \cdots \tag{8.86}$$

de donde obtenemos una jerarquía de ecuaciones

$$H_0\left(\frac{\partial W_0}{\partial\theta^1},\ldots,\frac{\partial W_0}{\partial\theta^n}\right) = E_0, \tag{8.87}$$

$$\sum_j \frac{\partial H_0}{\partial J_j}\frac{\partial W_1}{\partial \theta^j} + H_1\left(\frac{\partial W_0}{\partial \theta^1},\ldots,\frac{\partial W_0}{\partial \theta^n},\theta_1,\ldots\theta_n\right) = E_1, \tag{8.88}$$

$$\vdots \tag{8.89}$$

De la primera de estas ecuaciones obtenemos que

$$W_0 = \sum_j J_j\theta^j, \tag{8.90}$$

lo que nos indica que W vista como función generadora de transformaciones canónicas es, a orden cero, la generadora de la transformación idéntica, la que en coordenadas de acción y ángulo es en efecto solución de la ecuación de Hamilton Jacobi del sistema no perturbado.

A primer orden en ϵ obtenemos la segunda de las ecuaciones con la que obtendríamos W_1. Ahora recordamos que las variables θ son ángulos y por tanto W_1 y H_1 deben ser funciones periódicas de esos ángulos, lo que nos permite desarrollarlas en series de Fourier:

$$W_1 = \sum_{\bar{k}} A_{\bar{k}}(J)e^{i\bar{k}\cdot\bar{\theta}}, \tag{8.91}$$

y

$$H_1 = \sum_{\bar{k}} B_{\bar{k}}(J)e^{i\bar{k}\cdot\bar{\theta}}, \tag{8.92}$$

donde hemos denotado por $\bar{k}$ a una eneada cualquiera de números naturales y por $\bar{\theta}$ a $(\theta^1,\ldots,\theta^n)$. Hemos también utilizado el hecho de que

$$\frac{\partial W_0}{\partial \theta^j} = J_j. \tag{8.93}$$

La ecuación ((8.88)) quedará como

$$\sum_j \frac{\partial H_0}{\partial J_j}k_j\left[i\sum_{\bar{k}} A_{\bar{k}}(J)e^{i\bar{k}\cdot\bar{\theta}}\right] + \sum_{\bar{k}} B_{\bar{k}}(J)e^{i\bar{k}\cdot\bar{\theta}} = E_1 \tag{8.94}$$

o, recordando que $\partial H_0/\partial J_j$ es la frecuencia ω^j del sistema no perturbado

$$\sum_{\bar{k}}\left[\left(i\sum_j \omega^j k_j\right)A_{\bar{k}}(J) + B_{\bar{k}}(J)\right]e^{i\bar{k}\cdot\bar{\theta}} = E_1. \tag{8.95}$$

Esta ecuación nos indica que todos los términos de la suma deben ser cero, excepto el correspondiente a la eneada cero y por tanto

$$E_1 = B_0(J), \tag{8.96}$$

y

$$\left(i \sum_j \omega^j k_j \right) A_{\bar{k}}(J) + B_{\bar{k}}(J) = 0. \tag{8.97}$$

De esta última ecuación obtenemos

$$A_{\bar{k}}(J) = -i \frac{B_{\bar{k}}(J)}{\sum_j \omega^j k_j} \qquad \bar{k} \neq 0, \tag{8.98}$$

A_0 sería arbitraria por lo que la podemos tomar como cero. Al substituir estos valores de las $A_{\bar{k}}$ en ((8.91)) obtenemos la solución de la ecuación de Hamilton Jacobi a primer orden en ϵ, o como se acostumbra decir a primer orden en perturbaciones. Podríamos ahora continuar con la jerarquía de ecuaciones y encontrar la solución de la ecuación de Hamilton Jacobi al orden que se desee.

En principio esto nos proporciona la solución del sistema perturbado, al menos para ϵ suficientemente pequeña, pero hay una dificultad en ((8.98)), el término en el denominador

$$\sum_j \omega^j k_j \tag{8.99}$$

es una combinación de las ω con una de las eneadas de números naturales y como las ω son las frecuencias del sistema no perturbado, cuando son resonantes habrá valores de $\bar{k}$ para las que esa suma valga cero y entonces la serie ((8.91)) presenta términos infinitos y por tanto es divergente. Es más, en los casos en que los valores de las J correspondan a toros no resonantes del sistema no perturbado, habrá eneadas $\bar{k}$ para las que el término en el denominador de ((8.98)) se haga tan pequeño como se quiera y, por tanto, la serie ((8.91)) será también divergente. Concluimos pues que, si bien hemos obtenido una solución formal, esta es inútil, pues el método seguido presenta problemas básicos de convergencia.

Esta dificultad aparece en muchas otras formas en la teoría canónica de perturbaciones y se le conoce como el problema de los denominadores pequeños. La teoría de perturbaciones está plagada de estas divergencias, esto hace que sea sumamente complejo el desarrollarla y en especial utilizarla

para resolver cuestiones teóricas como el destino de los toros inconmensurables e incluso para encontrar en casos concretos el comportamiento de las trayectorias a tiempos grandes.

A pesar de ello se han encontrado métodos de perturbaciones útiles para atacar problemas concretos, en especial cuando no se pretende extender las soluciones a tiempos largos, esto es, tiempos para los que las resonancias hayan generado amplitudes grandes. Muchos son los sistemas en los que se han aplicado con éxito métodos de perturbaciones. No entraremos aquí al desarrollo de las técnicas basadas en la teoría de perturbaciones, cuestión que dejamos a los textos especializados, pero si señalamos que todos los intentos por resolver el problema de los denominadores pequeños u otros de naturaleza semejante, esto es, lograr series convergentes, no prosperaban, pues, como podemos ver es un problema íntimamente ligado con los acoplamientos resonantes que introduce la perturbación entre los grados de libertad. Poincaré mismo trato de darle solución y lo consideró como el problema fundamental de la mecánica clásica. Así el problema fundamental de la teoría de perturbaciones y con él el del destino de los toros inconmensurables, quedó sin solución.

A fines de los años cincuenta Kolmogorov ideo un método de muy rápida convergencia para el estudio de los sistemas hamiltonianos integrables perturbados. La idea básica es similar a la del método de Newton para encontrar las raíces de un polinomio o de una función en general. Escapa a este texto la exposición detallada de este método de Kolmogorov, el que desarrollado y aplicado por Arnold y Moser genera la demostración de un teorema hoy conocido por la iniciales KAM; indicaremos únicamente que extiende la idea iterativa del método de Newton a espacios de funciones, logrando con ello la convergencia buscada. Pasemos pues a ilustrar, utilizando el método de Newton para encontrar raíces de funciones, el proceso de convergencia rápida, para después pasar a enunciar el teorema de KAM, cuya demostración se sustenta en las ideas de Kolmogorov.

Pensemos en una función de una variable $f(x)$ y consideremos el cálculo de sus raíces, esto es la solución de la ecuación

$$f(x) = 0. \tag{8.100}$$

Imaginemos que tenemos un valor x_0 que es cercano a una raíz, entonces podemos plantearnos el problema de aproximar la raíz por un método de perturbaciones, esto es, partiendo de un desarrollo en serie para ella. Ahora, considerando que x_0 es cercano a la raíz, llamemos ϵ a $f(x_0)$ y busquemos la raíz como un desarrollo en serie en ϵ. Desarrollando en serie la función f

tendremos

$$f(x) = f_0 + f_0^{(1)}(x - x_0) + \frac{1}{2}f_0^{(2)}(x - x_0)^2 + \frac{1}{3!}f_0^{(3)}(x - x_0)^3 + \cdots \quad (8.101)$$

donde

$$f_0 = f(x_0) = \epsilon, \quad (8.102)$$

$$f_0^{(1)} = \left.\frac{df}{dx}\right|_{x_0}, \quad (8.103)$$

$$f_0^{(2)} = \left.\frac{d^2f}{dx^2}\right|_{x_0}, \quad (8.104)$$

$$\vdots \quad (8.105)$$

la ecuación para la raíz queda como

$$0 = \epsilon + f_0^{(1)}(x - x_0) + \frac{1}{2}f_0^{(2)}(x - x_0)^2 + \frac{1}{3!}f_0^{(3)}(x - x_0)^3 + \cdots \quad (8.106)$$

y por tanto tendremos que

$$-\frac{\epsilon}{f_0^{(1)}} = (x - x_0) + \frac{1}{2}\frac{f_0^{(2)}}{f_0^{(1)}}(x - x_0)^2 + \frac{1}{3!}\frac{f_0^{(3)}}{f_0^{(1)}}(x - x_0)^3 + \cdots . \quad (8.107)$$

Para encontrar x a partir de aquí usemos la teoría de perturbaciones, esto es, desarrollemos x en serie de ϵ

$$x = x_0 + x_1\epsilon + \frac{x_2}{2}\epsilon^2 + \frac{x_3}{3!}\epsilon^3 + \cdots , \quad (8.108)$$

al substituir en ((8.107)) obtenemos

$$-\frac{1}{f_0^{(1)}} = \left(x_1 + \frac{x_2}{2}\epsilon + \frac{x_3}{3!}\epsilon^2 + \cdots\right) + \frac{1}{2}\frac{f_0^{(2)}}{f_0^{(1)}}\epsilon\left(x_1 + \frac{x_2}{2}\epsilon + \frac{x_3}{3!}\epsilon^2 + \cdots\right)^2 + \quad (8.109)$$

$$+\frac{1}{6}\frac{f_0^{(3)}}{f_0^{(1)}}\epsilon^2\left(x_1 + \frac{x_2}{2}\epsilon + \frac{x_3}{3!}\epsilon^2 + \cdots\right)^3 + \cdots \quad (8.110)$$

al igualar los términos con la misma potencia de ϵ obtenemos los coeficientes de la serie ((8.108)), esto es

$$x_1 = -\frac{1}{f_0^{(1)}}, \quad (8.111)$$

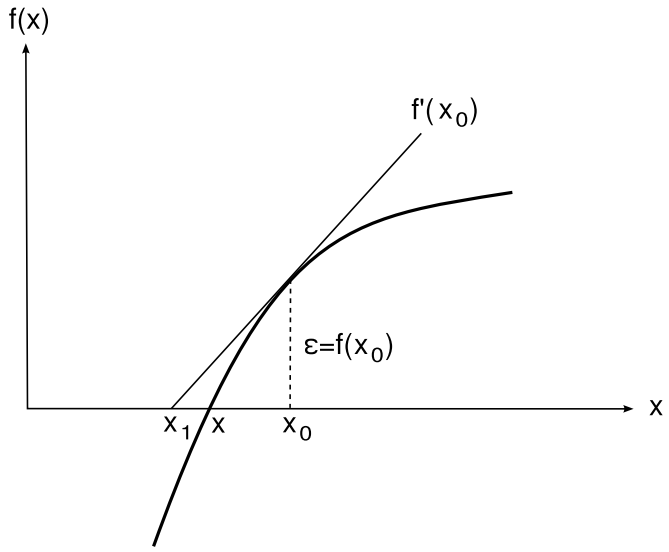


Figura 8.10: Primer término del método perturbativo para encontrar la raíz de una función.

$$x_2 = -x_1^2 \frac{f_0^{(2)}}{f_0^{(1)}} = -\frac{f_0^{(2)}}{\left(f_0^{(1)}\right)^3}, \tag{8.112}$$

$$x_3 = -\frac{3!}{2}\frac{f_0^{(2)}}{f_0^{(1)}} x_1 x_2 - \frac{f_0^{(3)}}{f_0^{(1)}} x_1^3 = -3\frac{f_0^{(2)}}{\left(f_0^{(1)}\right)^5} + \frac{f_0^{(3)}}{\left(f_0^{(1)}\right)^4}, \tag{8.113}$$

$$\vdots \tag{8.114}$$

y, por tanto

$$x = x_0 - \frac{1}{f_0^{(1)}}\epsilon - \frac{f_0^{(2)}}{2\left(f_0^{(1)}\right)^3}\epsilon^2 - \left[\frac{f_0^{(2)}}{2\left(f_0^{(1)}\right)^5} - \frac{f_0^{(3)}}{6\left(f_0^{(1)}\right)^4}\right]\epsilon^3 + \cdots . \tag{8.115}$$

Vemos como la convergencia de esta serie depende sustantivamente del valor de $f_0^{(1)}$ que está presente en todos los términos de las potencias de ϵ; desde luego esta forma de la serie presenta serios inconvenientes. Desde una perspectiva gráfica vemos en la figura (8.10) la aparición del primer término. Las siguientes aproximaciones dependen de las derivadas superiores de la función valuadas en x_0.

Veamos ahora el método de Newton para encontrar las raíces de la función $f(x)$.[6]

Tomamos una primera aproximación que coincide con la ya encontrada a primer orden, esto es, llamando X_1 a esta aproximación

$$X_1 = x_0 - \frac{f(x_0)}{f^{(1)}(x_0)}, \tag{8.116}$$

ahora, en vez de buscar un término de segundo orden sustentandonos en el valor de la función y sus derivadas en x_0, repitamos el procedimiento de encontrar una primera aproximación pero partiendo del punto X_1; llegaremos al punto X_2 dado por

$$X_2 = X_1 - \frac{f(X_1)}{f^{(1)}(X_1)}, \tag{8.117}$$

y así sucesivamente siguiendo un proceso de iteración

$$X_{n+1} = X_n - \frac{f(X_n)}{f^{(1)}(X_n)}. \tag{8.118}$$

Gráficamente podemos seguir el proceso como se indica en la figura (8.11).

Podemos ver que

$$X_{n+1} - X_n = -\frac{f(X_n)}{f^{(1)}(X_n)}, \tag{8.119}$$

por lo que la sucesión dada por la iteración converge a la raíz si la cantidad

$$\delta_{n+1} = X_{n+1} - X_n \tag{8.120}$$

tiende a cero. Ahora

$$\delta_{n+1} = -\frac{f(X_n)}{f^{(1)}(X_n)}, \tag{8.121}$$

pero

$$X_n = X_{n-1} - \frac{f(X_{n-1})}{f^{(1)}(X_{n-1})} \tag{8.122}$$

entonces

$$\delta_{n+1} = -\frac{f\left(X_{n-1} - \frac{f(X_{n-1})}{f^{(1)}(X_{n-1})}\right)}{f^{(1)}\left(X_{n-1} - \frac{f(X_{n-1})}{f^{(1)}(X_{n-1})}\right)} = -\frac{f(X_{n-1} - \delta_n)}{f^{(1)}(X_{n-1} - \delta_n)}, \tag{8.123}$$

[6] Newton desarrollo este método únicamente para polinomios en x, la extensión a una función arbitraria es posterior.

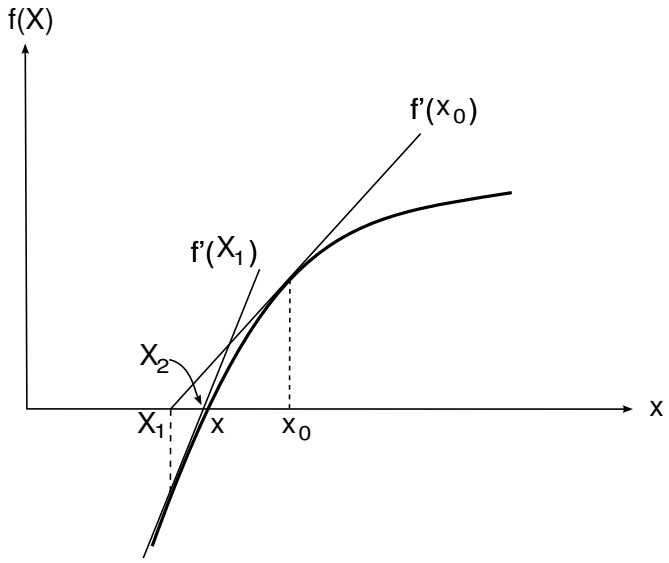


Figura 8.11: Dos iteraciones en el método de Newton para encontrar las raíces de una función.

por tanto

$$\delta_{n+1} = -\frac{f(X_{n-1})}{f^{(1)}(X_{n-1})} + \frac{f^{(1)}(X_{n-1})^2 - f(X_{n-1})f^{(2)}(X_{n-1})}{f^{(1)}(X_{n-1})^2}\delta_n + \cdots = \tag{8.124}$$

$$-\frac{f^{(2)}(X_{n-1})}{f^{(1)}(X_{n-1})}\delta_n^2 + \cdots . \tag{8.125}$$

Vemos entonces que la convergencia es exponencial, pues los términos de la sucesión de aproximaciones van como δ^{2^n}, esto es, si la sucesión converge, converge exponencialmente y el proceso de iteraciones será mucho más eficiente que el desarrollo directo en serie. La principal diferencia estriba en que, en la serie se utiliza, para acercanos a la raíz, la derivada de la función valuada siempre en el punto inicial, mientras que en el método de Newton se utiliza en cada paso la derivada de la función valuada en el nuevo punto, lo que potencia la convergencia cuando la hay.

Utilizando técnicas semejantes a la descrita se logra estudiar el problema de los toros no resonantes, evitando las dificultades de convergencia que presenta la teoría de perturbaciones clásica. El resultado es el famoso teorema KAM, de Kolmogorov, Arnold y Moser, que nos dice que:

Si tenemos un sistema hamiltoniano integrable, acotado y perturbado

$H = H_0 + \epsilon H_1$, con H suficientemente derivable y con H_0 no degenerada, esto es, tal que

$$\left| \frac{\partial^2 H_0}{\partial J_i \partial J_j} \right| \neq 0 \tag{8.126}$$

donde las J son las variables de acción para H_0, entonces para ϵ suficientemente pequeña, una fracción de los toros inconmensurables no se destruyen, únicamente se deforman. Esta fracción de toros ocupa una medida distinta de cero en el espacio fase, medida que tiende a uno cuando ϵ tiende a cero, pero que tiende a cero cundo ϵ tiende a un cierto valor crítico, esto es, para ϵ mayor que ese valor todos los toros se habrán roto; para valores intermedios de ϵ la medida se comparte entre los toros que permanecen y los espacios o intersticios que van dejando los toros rotos, intersticios en que se desarrollan trayectorias que no trazan toros. Naturalmente los toros que permanecen están integrados por trayectorias cuasiperiódicas, puesto que los compuestos por trayectorías periódicas son inestables ante la más mínima perturbación.

Es importante recordar que en el instante en que la perturbación deja de ser cero, todos los toros resonantes se rompen, por lo que el conjunto de toros que no se rompen forman un conjunto de cantor, que se forma, a partir del conjunto continuo de toros del sistema no perturbado, al romperse los toros resonantes que están intercalados en el continuo de todos ellos como los números racionales se intercalan ente los reales.

Así al actuar la perturbación con ϵ suficientemente pequeña nos queda un sistema cantoriano de toros e intersticios; sobre los toros que quedan tendremos trayectorias cuasiperiódicas; los intersticios son los espacios dejados por los toros rotos, los resonantes y los no resonantes que se hayan roto para ese valor de ϵ, en estos intersticios las trayectoria serán complejas y algunas de ellas estarán organizadas en la forma que describimos al estudiar el destino de los toros resonantes.

Usando el mapeo de Poincaré para el caso de dos grados de libertad la imagen esquemática que tendremos se muestra en la figura (8.12).

Los círculos están formados por las trazas que dejan, siguiendo el mapeo de Poincaré, las trayectorias cuasiperiódicas sobre los toros que no se han roto y los espacios representan los intersticios, en ellos hemos puesto únicamente las trazas que dejan las trayectorias periódicas aisladas que surgen al romperse los toros resonantes y parte de las trazas sobre las variedades estable e inestable en el caso de las inestables y los toros que se forman en torno a las estables, naturalmente en estos intersticios hay muchísimas más trazas de muchas otras trayectorias que no hemos dibujado, entre ellas las resultantes de toros inconmensurables que ya se han roto. Es importante recordar que

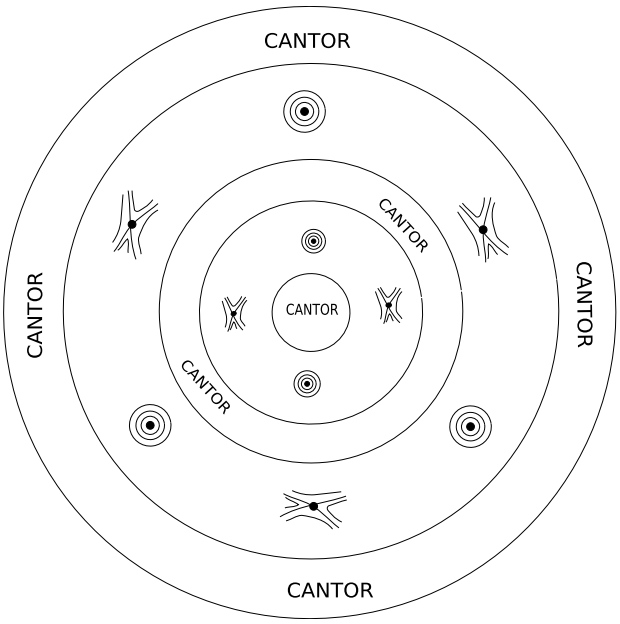


Figura 8.12: Los toros cuasiperiódicos que no se rompen se intercalan entre los espacios o intersticios dejados por los toros que si se han roto formando una estructura de Cantor.

los círculos forman una estructura cantoriana que no podemos reproducir con fidelidad en un dibujo, pero que implica que los círculos, y por tanto los toros, están aislados, a pesar de que hay un número no numerable de ellos y ocupan una medida finita; es más si ϵ es muy pequeño esa medida será casi uno, sin embargo hay intersticios tan cercanos a cualquier círculo como se quiera. A medida que ϵ crece los intersticios se van haciendo grandes, se rompen más y más toros, pero sin dejar de formar una estructura cantoriana, hasta que se alcanza el valor critico de ϵ en que se rompe el 'último' toro y toda la estructura se pierde.

En cuanto a los círculos que hemos trazado en torno a los puntos que representan las órbitas periódicas aisladas provenientes de la rotura de los toros resonantes, recordamos que se trata de toros que se enrollan varias veces antes de cerrarse y que están compuestos por trayectorias cuasiperiódicas, esto implica que están sujetos a las mismas condiciones que los otros, esto es, estos toros se irán rompiendo a medida que ϵ crezca y formarán una estructura cantoriana con intersticios que tienen propiedades similares a la estructura general. Naturalmente esto produce un sistema fractal en el que, al tomar una ampliación de cualquiera de las regiones que incluyen estos toros, se reproduce una estructura similar a la del conjunto más amplio, llegándose a repetir y repetir esto de manera interminable. Ver figura (8.13).

Esta es la estructura que queda al perturbar un sistema integrable y nos explica el resultado de los cálculos de FPU, simplemente la perturbación no alcanza a romper todos los toros, es más, la medida de los toros que permanecen es aun grande, por lo que las trayectorias son, en su mayoría cuasiperiódicas, razón por la cual después de un tiempo la energía se reconcentra en el modo normal en que la situaron las condiciones iniciales.[7]

Así encontramos que un importante número de trayectorias serán cuasiperiódicas si ϵ es suficientemente pequeña, pero excepto por el comportamiento en vecindades pequeñas de las trayectorias periódicas que resultan de la desintegración de los toros resonantes, no hemos encontrado el comportamiento de las trayectorias en los intersticios que se generan al irse rompiendo los toros, pasemos pues a estudiar esos casos, que la ir aumentando ϵ se van haciendo dominantes.

[7] El caso FPU tiene algunos problemas con la condición ((8.126)), pero el resultado del teorema de KAM es aplicable.

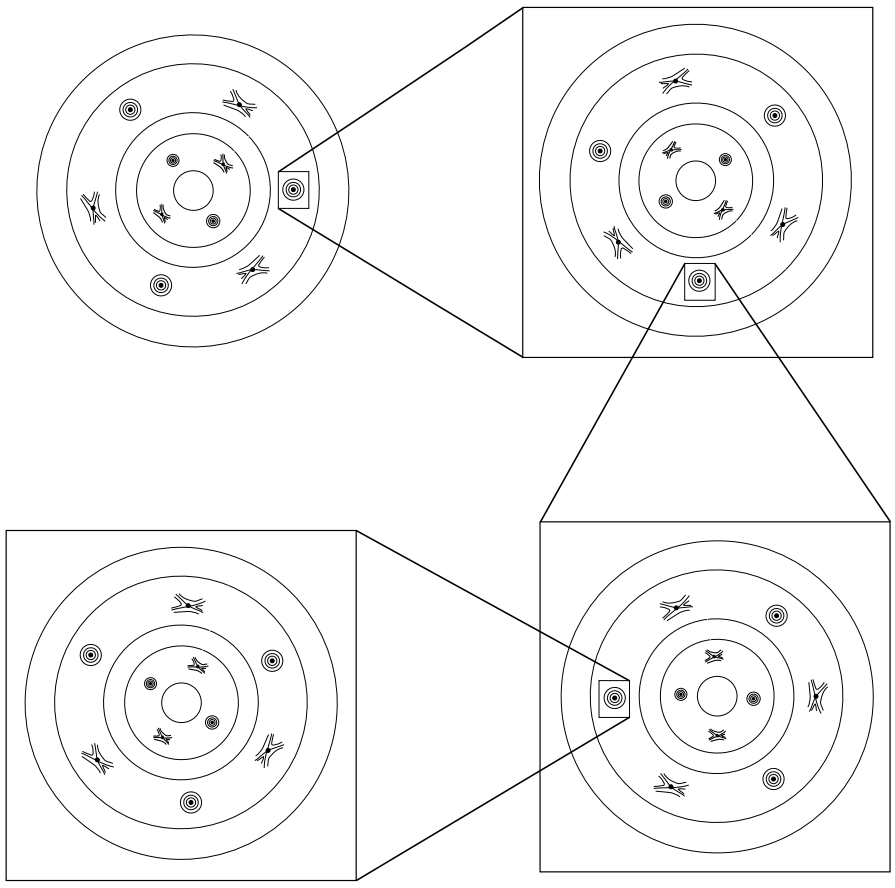


Figura 8.13: En una vecindad de un punto fijo estable se reproduce la misma estructura de toros cuasiperiódicos e intersticios. A su vez esto se vuelve a reproducir generándose una estructura fractal.

8.4. Caos hamiltoniano.

Para hacer el estudio del comportamiento de las trayectorias dentro de los intersticios que quedan entre los toros cuasiperiódicos, utilizaremos nuevamente el mapeo de Poincaré, sin olvidar que sus órbitas tendrán un comportamiento que refleja el de las trayectorias del sistema. Tomemos uno de los puntos que constituyen una órbita periódica hiperbólica inestable proveniente de la rotura de un toro resonante. Un punto del plano perteneciente a una órbita periódica resultante de la ruptura, tendrá una variedad estable y una inestable compuestas por las órbitas que tienden asintóticamente al punto o que divergen asintóticamente de él; ahora estas variedades, que en el caso de dos grados de libertad son curvas, no pueden intersectar ninguno de los toros que permanecen ni terminar en la nada, por lo que o se extienden indefinidamente o tienen algún tipo de comportamiento asintótico por el otro extremo. En el segundo caso tendríamos situaciones en que se une suavemente una variedad estable con una inestable como se muestra en la figura (8.14). Se puede demostrar que estas uniones son inestables, esto es, cualquier pequeña perturbación o cambio en ϵ, las destruye, por lo que genéricamente hablando no existen, esto es, el comportamiento genérico es el extenderse indefinidamente, naturalmente sin intersectar ninguno de los toros.

Esto conduce a que las variedades estables e inestables se extiendan y se vayan enredando de maneras complejas pudiendo ocurrir que jamas se intersectarán entre ellas o que si lo hicieran. Se puede hacer ver que el caso en que no se producen intersecciones es estructuralmente inestable, por lo que concluimos que genéricamente se producen intersecciones que, naturalmente, deben ser entre una variedad estable y una inestable[8], el punto de intersección pertenece a una órbita del mapeo que tiende asintóticamente a uno de los puntos y que sale asintóticamente de otro o del mismo, como se ilustra en la figura (8.15); en el primer caso se dice que se ha formado un intersección heteroclínica y en el segundo una intersección homoclínica.

Debemos recordar que la ruptura de un toro resonante produce una trayectoria periódica que, en el terreno del mapeo de Poincaré, vemos como una órbita consistente de un cierto número de puntos que se van mapeando entre si en una secuencia periódica; entonces, el mapeo de Poincaré compuesto tantas veces como el período, presentará una colección de puntos fijos, uno por cada uno de los puntos que constituyen la órbita periódica del mapeo

[8]Debemos recordar que se trata de las variedades estables e inestables y no de trayectorias de una ecuación diferencial, las que desde luego no podrían intersectarse.

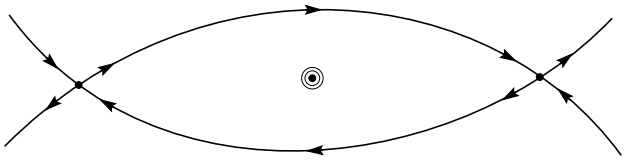


Figura 8.14: Las variedades estables e inestables de dos puntos fijos hiperbólicos se unen suavemente. Esta unión es estructuralmente inestable.

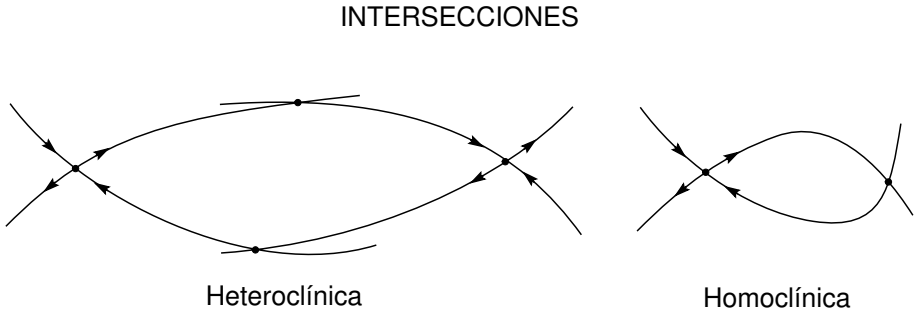


Figura 8.15: Al intersectar+se la variedad estable con la inestable de dos puntos fijos distintos se forma una intersección heteroclínica, si trata de las variedades estable e inestable del mismo punto, la intersección es homoclínica.

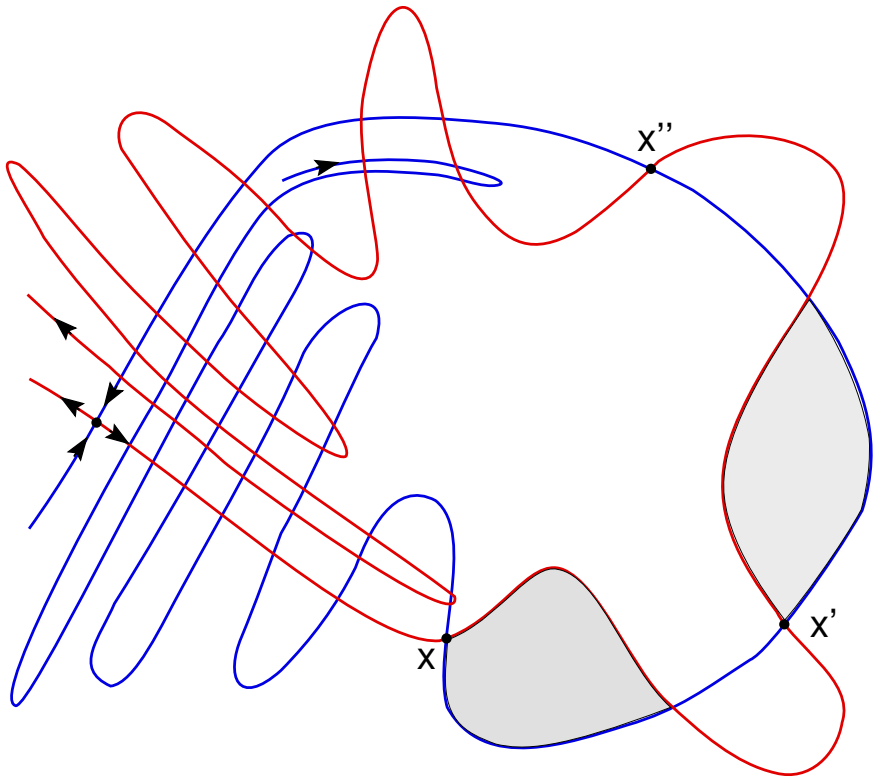


Figura 8.16: El enredo homoclínico.

simple. Así, con la excepción de las intersecciones que pudieran ocurrir entre alguna de las variedades de una órbita con las de otra órbita distinta, el caso más frecuente es el segundo aunque para el mapeo compuesto aparezca como heteroclínico.

Ahora al haber una intersección homoclínica, necesariamente habrá muchas, veamos. El punto de intersección, que hemos denotado por x en la figura (8.16), pertenece a dos variedades, una de ellas estable y la otra inestable, por lo que se mapeará en un punto x' que tendrá que pertenecer a cada una de esas dos variedades, esto es, conformará una nueva intersección, que deberá estar en la dirección correcta en ambas variedades. A su vez x' se mapeará en x'' que tendrá que ser una nueva intersección y así sucesivamente, como se ilustra en la figura. Naturalmente el punto de intersección se irá hacercando asintóticamente al punto fijo, por lo que las distancias entre ellos se irán acortando y tenderán a cero.

Esto último produce un efecto muy importante. Observemos el área som-

breada en la figura, al aplicar el mapeo a los puntos de su interior tendrán que mapearse en puntos del interior de la otra área sombrada, de lo contrario la trayectoria correspondiente intersectaría alguna de las variedades invariantes, lo que no es posible. Ahora este mapeo proviene de un sistema hamiltoniano que, por el teorema de Liouville, preserva el volumen en el espacio fase, esto implica que el mapeo de Poincaré preserva área, por tanto, el área encerrada en las regiones sombreadas es la misma, pero la distancia entre las iteraciones sucesivas tiende a cero, lo que obliga a un crecimiento de la región en la otra dirección, produciéndose lo ilustrado en la figura.

Por otro lado si aplicamos el mapeo inverso a x, esto es, vemos de que punto de la órbita proviene x, nos damos cuenta que ocurre lo mismo, pues x no puede ser más que la imagen de un punto que pertenezca a ambas variedades. Al ir iterando en sentido inverso obtenemos algo totalmente similar a lo anterior donde el papel de la variedad estable lo tiene ahora la inestable y viceversa. Así obtenemos una configuración en la que necesariamente se producen muchas más intersecciones al cruzarse los lóbulos que conforman ambas variedades. En adición a esto también observamos que, dado que estas variedades no se pueden intersectar a si mismas, los lóbulos quedan encerrados, pero deben crecer en longitud por lo que se deben doblar y doblar, lo que a su vez genera más intersecciones y así sucesivamente. Podemos observar la enorme complejidad que se produce. En palabras del propio Poincaré ´Las intersecciones forman un especie de red, trama o tejido con bucles infinitamente apretados; ninguna de las dos curvas (las variedades) podrán alguna vez intersectarse a si mismas y deben doblarse en una forma tan compleja que intersectarán a todos los bucles de la red un número infinito de veces.

´Uno queda apabullado por la complejidad de esta figura que ni siquiera me atrevo a dibujar. Nada puede darnos una mejor idea de la complejidad del problema de tres cuerpos...´

La complejidad de esto se ve aún mayor cuando recordamos que estamos viendo únicamente los puntos en que las trayectorias del sistema mecánico intersectan a la sección de Poincaré, que estamos viendo como resultado de aplicar el mapeo compuesto, que esto ocurre en todos los intersticios que se forman en el conjunto de cantor de toros cuasiperiódicos y que, en los entornos de cada órbita periódica estable tendremos un cantor de toros en cuyos intersticios volverá a parecer lo mismo en una reproducción infinita del esquema.

En el caso de tres o más grados de libertad se reproduce el esquema básico que hemos establecido para dos. Sin embargo, en esos casos surgen complicaciones adicionales producto del hecho de que los intersticios que se forman entre los toros ya no serán regiones acotadas del espacio fase. Esto

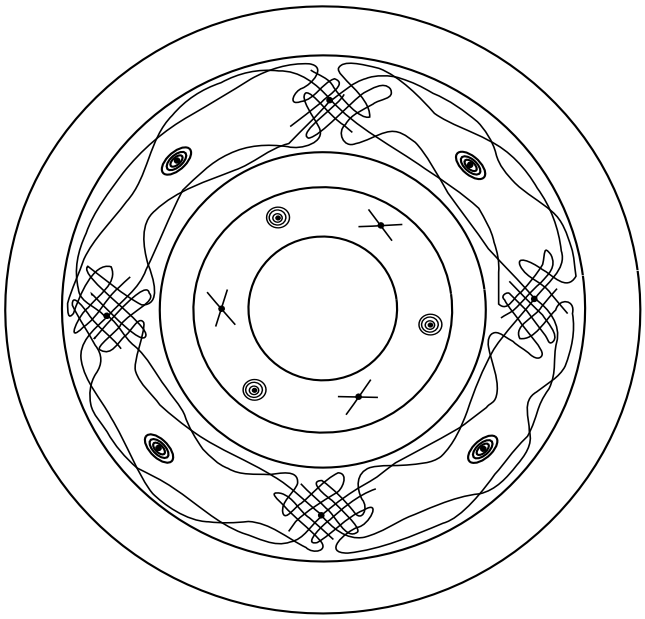


Figura 8.17: En los intersticios que la ruptura de los toros resonantes dejan se producen marañas homoclínicas, sumamente complejas. Por razones de claridad en la figura únicamente hemos dibujado las marañas en uno de los intersticios, pero hay que recordar que ocurren en todos.

se debe a que los toros concéntricos de dimensión mayor que dos, inmersos en espacios del doble de su dimensión, no acotan el espacio entre ellos como ocurre con los toros de dos dimensiones en un espacio de cuatro. Esto último da origen al hecho de que las trayectorias en los intersticios puedan extenderse por regiones abiertas de la variedad de energía constante, fenómeno conocido como difusión de Arnold. Otra cuestión que complica el caso de más de dos grados de libertad es el hecho de que las variedades estables e inestables son variedades de dimensión mayor que uno.

Pasemos ahora a ver, como ejemplo ilustrativo de las propiedades de las trayectorias que pueden ocurrir en los sistemas no integrables, un caso específico del problema restringido de tres cuerpos estudiado por Sitnikov y Alekseev.

Pensemos en un sistema estelar binario consistente de dos estrellas de la misma masa M; las dos estrellas trazarán sobre un plano órbitas keplerianas elípticas colocadas simétricamente en torno al centro de masas del sistema como se ilustra en la figura (8.18). Consideremos ahora un pequeño planeta de masa m, despreciable comparada con la masa de las estrellas, y que fue colocado inicialmente sobre la recta perpendicular al plano de las órbitas de las estrellas que pasa por el centro de masas de éstas y con una velocidad inicial a lo largo de esa misma recta. Por masa despreciable entenderemos aquella que por su pequeña magnitud nos permite despreciar totalmente el efecto del planeta sobre el movimiento de las estrellas, por lo que podemos estudiar el sistema considerando que las estrellas se mueven en órbitas keplerianas; a este tipo de aproximaciones se les conoce con el nombre general de problemas restringidos de tres cuerpos, naturalmente la fuerza de gravedad de las estrellas si actuará sobre el planeta.

De la figura (8.18) vemos que, por simetría, la fuerza total sobre el planeta actúa siempre a lo largo de la recta perpendicular al plano que pasa por el centro de masas y dadas las condiciones iniciales especificadas para el planeta la posición de éste permanecerá siempre sobre esta recta. En general el planeta presentará un movimiento oscilatorio en torno al centro de masas de las estrellas.

La magnitud de la fuerza sobre el planeta será

$$F = \frac{2GMmx}{[x^2 + r(t)^2]^{\frac{3}{2}}} \tag{8.127}$$

donde x es la posición del planeta y $r(t)$ la distancia de las estrellas a su centro de masas dada por la ec**** de la solución al problema de Kepler.

Si las órbitas de las estrellas fuesen circulares, esto es para excentricidad cero, r sería una constante R_0 y el movimiento del planeta resultaría

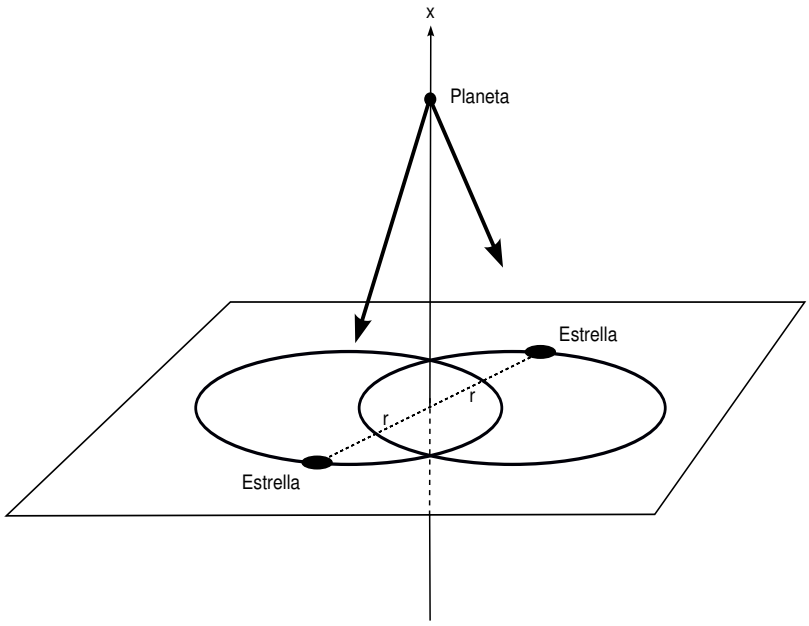


Figura 8.18: El sistema de Sitnikov y Alekseev.

integrable, por lo que el planeta realizaría oscilaciones periódicas o tendería monótonamente al infinito dependiendo de las condiciones iniciales. Para excentricidades ϵ pequeñas podemos desarrollar en términos de ϵ y quedarnos a primer orden, obtendríamos para $r(t)$ una aproximación de la forma

$$r(t) = R_0(1 - \epsilon A cos(\omega t)) + o(\epsilon^2) \tag{8.128}$$

que al substituir en la expresión de la fuerza nos proporcionaría una ecuación de movimiento a primer orden. La integración de esta ecuación de movimiento no autónoma no es posible hacerla, sin embargo se tiene algunos resultados importantes y sorprendentes sobre este sistema, veamos esto.

Tomemos una sucesión cualquiera de números reales $\cdots s_{-2}s_{-1}s_0s_1s_2\cdots$ finita o infinita por un lado o por los dos. Ahora tomemos la sucesión de intervalos de tiempo entre dos pasos consecutivos del planeta por el centro de masas de las estrellas, digamos $\cdots \tau_{-2}\tau_{-1}\tau_0\tau_1\tau_2\cdots$, donde una sucesión finita por la izquierda significa que el planeta venía de infinito y es capturado por el sistema y una finita por la derecha indica que el planeta escapa a infinito después de su último paso por el centro de masas. Se ha demostrado que dada la sucesión arbitraria de números reales, existen condiciones iniciales para el planeta, esto es una posición x_0 y una velocidad $\dot{x}_0$, tales que la sucesión de

intervalos de tiempo entre pasos sucesivos del planeta por el centro de masas reproduce exactamente la sucesión de números reales dada. Para aclarar el significado de esto veamos algunos ejemplos:

Si la sucesión de números reales es $\cdots 7.8, 15.3, 7.8, 15.3, 7.8, 15.3, \cdots$ esto es, una sucesión periódica, entonces con las condiciones iniciales adecuadas el planeta tendrá una órbita periódica con exactamente esos períodos.

Si la sucesión es $\cdots 7.8, 15.3, 7.8, 15.3, 7.8, 15.3$, nótese que es casi la misma del ejemplo anterior, excepto que es finita por la derecha, dadas las condiciones iniciales adecuadas el planeta estará en una órbita que después de pasar multitud de veces por el centro de masas en los tiempos indicados, escapa al infinito y no vuelve a pasar por el centro de masas. Nótese que esto implica que si se ha observado por muchos ciclos al planeta y se tienen como datos la suscesión de intervalos de tiempo entre pasos por el centro de masas, no es posible calcular el tiempo del siguiente paso, pues las condiciones pudieran ser las que dan origen a la órbita periódica, las que determinan que en un momento dado escapa a infinito o cualquier otra que habiendo recorrido la del ejemplo inicial continúe de otra manera, por ejemplo $\cdots 7.8, 15.3, 7.8, 15.3, 7.8, 15.3, 354.2, 27.1, 3.14159, \cdots$.

Si la sucesión es totalmente aleatoria, hay unas condiciones iniciales tales que los intervalos de tiempo entre pasos sucesivos del planeta la reproducen. Ahora al observar al planeta y tomar esos tiempos, no podemos distinguir si son producto del azar o, como es el caso, producto del sistema, totalmente determinista, consistente del planeta y las estrellas. Es más el teorema nos enseña que hay una correspondencia entre el conjunto de las parejas de números reales (posición y velocidad iniciales) y el de todas las sucesiones de números reales, ahora en el conjunto de sucesiones tendremos una infinidad de ellas periódicas, por lo que hay una infinidad de movimientos periódicos del planeta; pero el subconjunto de sucesiones aleatorias tiene medida uno, lo que implica que en la absoluta mayoría de los casos el planeta presentará un comportamiento indistinguible del aleatorio al que llamamos caótico; en adición a esto tenemos que el conjunto de sucesiones periódicas es denso aunque tenga medida cero, esto es en una vecindad de cualquier movimiento del planeta hay uno periódico; esto implica también que pequeños cambios en las condiciones iniciales pueden generar enormes diferencias en el comportamiento del planeta.

Veamos este mismo fenómeno en un caso mucho más simple. El ejemplo que usaremos se refiere a las iteraciones de un mapeo del círculo en si mismo. En realidad es un caso que no puede provenir de la mecánica clásica pues no preserva volumen, que en este caso de una dimensión se trata de la distancia; lo usamos porque muestra con nitidez el origen del comportamiento caótico

en los sistemas dinámicos en general y ,de manera similar, en los sistemas hamiltonianos.

Denotemos por θ a una coordenada angular para el círculo, por conveniencia parametrizaremos el círculo de 0 a 1 en vez de 0 a 2π. en estas coordenadas el mapeo es

$$\theta' = 2\theta \qquad (\text{mod } 1), \tag{8.129}$$

esto es, duplicamos el ángulo y colocamos el resultado sobre el círculo.

Notamos que el mapeo tiene un punto fijo $\theta = 0$, tiene también órbitas que después de un número finito de iteraciones aterrizan en el punto fijo, por ejemplo la que inicia en $\theta_0 = 1/4$, en la primera iteración tendremos $\theta_1 = 1/2$ y en la siguiente $\theta_2 = 0$, a partir de aquí todas serán cero ya que la órbita llegó al punto fijo. Hay también órbitas periódicas distintas al punto fijo, por ejemplo si $\theta_0 = 1/3$, $\theta_1 = 2/3$ y $\theta_3 = 4/3 \qquad (\text{mod } 1) = 1/3$, a partir de aquí se repetirá lo mismo, esto es una órbita de período dos. Hay órbitas que después de un cierto número de iteraciones convergen a una periódica, por ejemplo es fácil ver que si $\theta_0 = 19/24$ después de tres pasos se llega a la órbita de período dos anterior.

Este sistema tiene propiedades muy parecidas a las del caso de tres cuerpos. Dividamos el círculo en dos partes, $\theta \in [0, 0.5)$ y $\theta \in [0.5, 1)$, si al ir iterando el mapeo θ cae en el primer intervalo le asociamos la letra I o cualquier otro símbolo, si cae en el segundo le asociamos la letra D o cualquier otro símbolo distinto al primero; ahora construimos una sucesión de estos dos símbolos según vaya cayendo la órbita en uno u otro intervalo al irse desarrollando, a la secuencia que se obtiene se le llama un itinerario del sistema dinámico. Podemos demostrar el siguiente resultado sorprendente: Dada cualquier sucesión, finita o infinita, de los símbolos I y D, existe una condición inicial θ_0 tal que el itinerario de la órbita que comienza en θ_0 reproduce la sucesión dada. Esto es totalmente similar al caso del planeta, tenemos sucesiones periódicas de cualquier período, sucesiones que comienzan igual y luego se separan, sucesiones que después de un cierto número de pasos se hacen periódicas, sucesiones que comienzan periódicas y después dejan de serlo y sucesiones totalmente aleatorias; todas ellas pueden ser generadas por el sistema dinámico, basta comenzar en la condición inicial correspondiente. Naturalmente todos los problemas de predictivilidad que se tiene en el caso del planeta se reproducen aquí, la única diferencia está en que en este caso el mapeo no es invertible. Pasemos pues a demostrar que el sistema presenta el comportamiento descrito.

Expresemos la variable θ en sistema binario, entoneces $\theta = 0.a_1a_2a_3\ldots$, donde las a son cero o uno. Ahora, en sistema binario, multiplicar por dos es

lo mismo que recorrer el punto un lugar a la derecha, así $2\theta = a_1.a_2a_3\ldots$. Por otro lado la operación de tomar el módulo uno equivale a poner cero en lo que queda a la izquierda del punto, así tendremos que el mapeo en sistema binario queda como $2\theta \pmod 1 = 0.a_2a_3\ldots$, esto es, el mapeo equivale la operación recorrer el punto un lugar a la derecha e igualar a cero lo que quede a la izquierda del punto. Queda claro que dada una sucesión de ceros y unos, o cualquier otros par de símbolos, podemos generar una condición inicial θ_0 simplemente construyendo el número real correspondiente en sistema binario, al aplicar el mapeo a esta condición inicial e iterar se reproducirá la sucesión en la cifra que quede a la derecha del punto. Recordamos que cuando esta cifra es cero la variable o coordenada $\theta \in [0, 0.5)$ y de ser uno $\theta \in [0.5, 1)$.

Lo que ocurre en el mapeo del círculo que acabamos de estudiar puede generalizarse a una clase de sistemas dinámicos que llamamos corrimientos de Bernoulli, veamos esto. Tomemos un conjunto finito de símbolos $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ y construyamos el conjunto de todas las posibles secuencias de estos símbolos

$$\sum\nolimits_k^+ = \{S | S = s_0 s_1 \cdots , \quad s_i \in \alpha\}. \tag{8.130}$$

En este conjunto de secuencias se puede definir una métrica por medio de

$$d(S, S') = \sum_{i=0}^{\infty} \frac{\delta(s_i, s_i')}{2^i}, \tag{8.131}$$

donde donde $\delta(s_i, s_i')$ es cero si $s_i = s_i'$ y uno en caso contrario. Es fácil ver que esta suma siempre converge y que establece una métrica. Sustentada en esta métrica se puede construir una topología y hacer del conjunto $\sum_k^+$ un espacio topológico métrico.

Definamos ahora el mapeo $\sigma : \sum_k^+ \to \sum_k^+$ por medio de

$$\sigma(s_0 s_1 s_2 \cdots) = s_1 s_2 \cdots . \tag{8.132}$$

Este mapeo se conoce como un corrimiento de Bernoulli y es fácil hacer ver que se trata de un mapeo continuo en la topología que hemos definido para $\sum_k^+$.

El mapeo σ establece un sistema dinámico en el espacio $\sum_k^+$ en el que se pueden ver las propiedades que hemos estado asignando a un sistema caótico: Existen órbitas periódicas de cualquier período, las que tienen como condición inicial una secuencia periódica. Existe al menos una órbita densa, la que tiene como condición inicial una secuencia construida como la unión de secuencias que aproximan a cualquier otra. Hay alta sensibilidad a condiciones iniciales, dos secuencias que comienzan igual y después de n términos

se separan, constituyen condiciones iniciales cercanas que después de iterar n veces divergen.

En general diremos que un sistema dinámico definido en un espacio métrico $f : M \to M$ (un flujo $\psi_t : M \to M$ en el caso continuo) es caótico si tiene dos características:

1.- Alta sensibilidad a condiciones iniciales.

2.- Es topológicamente transitivo.[9]

Por alta sensibilidad a condiciones iniciales entendemos que existe una cantidad $r > 0$ tal que dado cualquier punto $x \in M$ y un cantidad arbitraria $\epsilon > 0$, existe un punto $y \in M$ tal que $d(x, y) < \epsilon$ y existe un entero $k > 0$ (o un real $t > 0$) para el cual $d(f^{(k)}(x), f^{(k)}(y)) \geq r$ $(d(\psi_t(x), \psi_t(y)) \geq r)$.

Diremos que el sistema dinámico es topológicamente transitivo si dados dos subconjuntos abiertos $U, V \subset M$ existe un entero $k > 0$ (un real $t > 0$) tales que $f^{(k)}(U) \cap V \neq \phi$ $(\psi_t(U) \cap V \neq \phi)$. Esto implica que M no puede separarse en dos o más subconjuntos invariantes ante el sistema dinámico, pues una pequeña vecindad pasará al irse desarrollando la dinámica tan cerca de cualquier punto como se quiera.

Es fácil ver que los corrimientos de Bernoulli tienen estas dos propiedades, esto es constituyen sistemas dinámico caóticos.

Podríamos pensar que estos corrimientos de Bernoulli son casos excepcionales entre los sistemas dinámicos, en realidad esto no es así. Muchas veces ocurre que, en un sentido muy amplio, lo que ocurre en un sistema dinámico definido por un mapeo en un espacio, es lo mismo que le sucede a otro mapeo en otro espacio, y que el estudio de las propiedades de uno de ellos es más sencillo que el estudio correspondiente en el otro. Veamos esto.

Consideremos dos mapeos $M : A \to A$ y $N : B \to B$ y supongamos que existe un homeomorfismo entre A y B $h : A \to B$, diremos que el homeomorfismo h establece una conjugación topológica entre los dos mapeos si

$$h \circ M = N \circ h \tag{8.133}$$

Ahora si $x_0, x_1, x_2, \ldots$ es una órbita de M tendremos que

$$x_{i+1} = M(x_i) \tag{8.134}$$

y aplicando ((8.133)) tendremos

$$h(x_{i+1}) = N(h(x_i)), \tag{8.135}$$

[9] Algunos autores agregan una tercera característica, la existencia de un conjunto denso de órbitas (trayectorias en el caso continuo) periódicas. Para muchos sistemas esta característica está presente y en muchos casos es consecuencia de o implica a las otras dos.

lo que implica que $h(x_0), h(x_1), h(x_2), \ldots$ es una órbita de N; esto es, cuando dos sistemas son conjugados topológicos, el homeomorfismo mapea órbitas en órbitas y los dos sistemas tendrán la misma dinámica.

Es frecuente en el estudio de los sistemas dinámicos utilizar esta idea de conjugación topológica. Se busca la presencia de algún subconjunto invariante del sistema, esto es, un subconjunto del espacio fase tal que toda trayectoria que pase por alguno de sus puntos permanece en él, si sobre este subconjunto se pueda establecer una conjugación topológica con algún otro sistema ya conocido, muy especialmente con un corrimiento de Bernoulli, el sistema dinámico tendría las propiedades del sistema conocido y, en caso de que éste sea un corrimiento, pues las del corrimiento. Veamos un ejemplo de esto.

Uno de los sistemas dinámicos más estudiados y que se ha hecho paradigmático tanto en la investigación como en la enseñanza de estas cuestiones, es el llamado mapeo logístico. Este es el mapeo de la recta real en si misma dado por

$$x_{i+1} = \mu x_i(1 - x_i), \tag{8.136}$$

que denotaremos por L_μ. Naturalmente este mapeo no puede provenir de la mecánica clásica ya que no preserva volumen y no es invertible, pero lo utilizaremos como ejemplo ilustrativo.

Los puntos fijos del mapeo se determinan por la ecuación

$$x = \mu x(1 - x), \tag{8.137}$$

cuyas soluciones son

$$x = 0 \qquad \text{y} \qquad x = 1 - \frac{1}{\mu}. \tag{8.138}$$

La estabilidad de estos puntos fijos u órbitas periódicas de período uno, está determinada por el valor de la derivada del mapeo en esos puntos, si este valor está entre uno y menos uno el punto fijo será estable e inestable en caso contrario. La derivada es

$$\mu(1 - 2x) \tag{8.139}$$

que en los punto fijos toma los valores μ y $2 - \mu$, por lo que el origen será estable si $-1 < \mu < 1$ y el punto $1 - 1/\mu$ lo será si $1 < \mu < 3$. Vemos que para valores $\mu > 1$ las órbitas que comienzan en valores de x negativos tienden asintóticamente a menos infinito y si comienzan en valores mayores que uno, despúes de un paso adquieren valor negativo, por o que a continuación tienden asintóticamente a menos infinito; así todas las órbitas que no

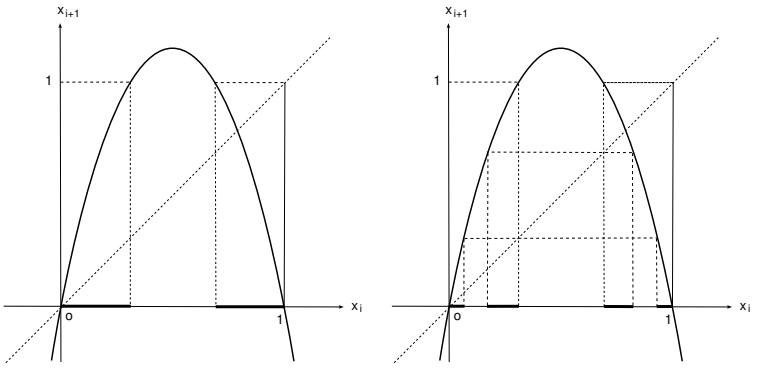


Figura 8.19: El mapeo logístico para $\mu > 2+\sqrt{5}$. En la gráfica de la izquierda se muestra el intervalo central que en la primera iteración se mapea a puntos con $x > 1$, y se han destacado con lineas gruesas los dos intervalos que se mapean dentro del intervalo $[0, 1]$. En la gráfica de la derecha se han destacado con lineas gruesas los cuatro intervalos que se mapean dentro del intervalo $[0, 1]$ en la segunda iteración.

comienzan en el intervalo $[0, 1]$ tienden a menos infinito. Las que comienzan en el intervalo $[0, 1]$ permanecerán siempre en ese intervalo si $\mu \leq 4$, esto es, si $0 \leq \mu \leq 4$ el intervalo I es invariante ante el mapeo. Este sistema dinámico ha sido muy estudiado para este rango de valores del parámetro μ y presenta una dinámica riquísima con la presencia de órbitas periódicas atractoras, el fenómeno de la bifurcación por doblamiento del período y otros muchos de interés, pero para los objetivos de ilustración que nos hemos puesto resulta más interesante analizar los casos en que $\mu > 2 + \sqrt{5}$.

Para $\mu > 2+\sqrt{5}$ el mapeo logístico se representa en la gráfica de la figura (8.19). Podemos observar que el intervalo I ya no es invariante, pues en la parte central se obtienen valores mayores que uno, por lo que los puntos del intervalo dentro de esa región central se mapean a puntos mayores que uno y posteriormente se irán asintóticamente a menos infinito. Por otro lado los puntos que quedan, ya sea en el intervalo que queda a la izquierda del central o su derecha, se mapean en puntos del propio intervalo I.

Ahora bien de entre los puntos de esos dos intervalos algunos se mapearán en el intervalo central, esos al continuar iterando abandonan el intervalo I y tenderán a menos infinito. El resto conforman cuatro pequeños intervalos, dos de ellos incluidos en el intervalo de la izquierda y los otros dos en el de la derecha, los puntos en esos cuatro intervalos se mapearán dentro de I en la segunda iteración. Esto continua una y otra vez a medida que iteramos, en

cada ocasión veremos que hay puntos cuya órbita abandona el intervalo I y que conforman intervalos incluidos en la parte central de todos los intervalos que quedaron en el paso anterior. Así al ir en cada paso quitando del intervalo I los puntos que lo abandonan, irán quedando aquellos cuya órbita queda incluida en I, después de un número infinito de pasos habremos quitado todos los que a fin de cuentas tienden a menos infinito y nos habremos quedado con los que tienen su órbita totalmente incluida en I. Como se van quitado intervalos abiertos nos quedará un conjunto no vacío de puntos cuyas órbitas permanecen siempre en I y que no pueden ser alcanzados más que por órbitas que provienen de puntos del propio I, esto es, se trata de un subconjunto invariante del sistema dinámico. La restricción del mapeo logístico a ese conjunto será entonces un sistema dinámico per se. Como podemos ver por construcción, este conjunto invariante, que se suele denotar con la letra Λ, es un conjunto de Cantor, es no numerable, no contiene ningún intervalo por pequeño que éste sea, cualquier intervalo que contenga un punto de él contendrá una infinidad y tiene medida cero, es interesante que siendo Λ un objeto tan complejo podamos sustentar en él todo un sistema dinámico.

Pasemos a estudiar la dinámica del mapeo logístico sobre el subconjunto invariante Λ. En primer término notamos que a Λ puede hacerse un espacio métrico, y por tanto topológico, simplemente utilizando la restricción a él de la métrica común en el intervalo I. Tomemos el intervalo $[0,1]$ y dividámosle en dos partes $I_- = [0, 1/2]$ y $I_+ = (1/2, 1]$. Ahora podemos asociar a cada órbita sobre Λ un itinerario compuesto de la sucesión de dos símbolos, digamos 0 o 1, según los puntos de la órbita vayan cayendo en los intervalos I_- o I_+. Así cada punto $x_0 \in \Lambda$ tendrá asociado un itinerario $S = s_0 s_1 s_2 \cdots$, donde s_0 será 0 o 1 según x_0 esté en I_- o en I_+, s_1 será 0 o 1 según $x_1 = F(x_0)$ esté en I_- o en I_+ y así sucesivamente.

Esta asociación de puntos de Λ con su itinerario se puede ver también como asociación de los puntos de Λ con los del espacio de secuencias de dos símbolos $\sum_2^+$. Esta asociación resulta ser un homeomorfismo $h : \Lambda \to \sum_2^+$ entre el conjunto de cantor invariante Λ y $\sum_2^+$, pasemos a demostrar esto.

En primer lugar demostraremos que el mapeo h es biyectivo. Supongamos que existen dos puntos distintos x y y de Λ tales que $h(x) = h(y)$, entonces x y y tendrían el mismo itinerario y ambos estarán en I_- o en I_+, por tanto el intervalo $I_{xy} = [x, y]$ se mapeará al aplicar el mapeo logístico en un intervalo $L_\mu(I_{xy})$ que estará totalmente contenido en I_- o en I_+, ya que $L_\mu(x)$ y $L_\mu(y)$ están en el mismo intervalo y el mapeo logístico es monótono en cada uno de los intervalos I_- e I_+. Naturalmente ocurrirá lo mismo al aplicar nuevamente el mapeo al intervalo resultante, esto implica que $L_\mu^{(n)}(I_{xy})$ será

un intervalo totalmente contenido en I_- o en I_+ para toda n. Entonces todos los puntos del intervalo $[x, y]$ tienen el mismo itinerario que x y y, y, como x y y son puntos de Λ, $[x, y] \subset \Lambda$, pero Λ es un conjunto de Cantor que no puede contener intervalo alguno. Hemos llegado a una contradicción por lo que x y y no pueden ser distintos, esto es h es inyectivo. Tomemos ahora un punto $S = s_0 s_1 s_2 \cdots$ de $\sum_2^+$ y busquemos un punto $x \in \Lambda$ tal que $h(x) = S$. Definamos una colección de mapeos por medio de

$$I_{s_0 s_1 \cdots s_n} = \{x \in I | x \in I_{s_0}, L_\mu(x) \in I_{s_1}, L_\mu^{(2)}(x) \in I_{s_2}, \ldots, L_\mu^{(n)}(x) \in I_{s_n}\} \tag{8.140}$$

$$= I_{s_0} \cap L_\mu^{(-1)}(I_{s_1}) \cap L_\mu^{(-2)}(I_{s_2}) \cap \cdots \cap L_\mu^{(-n)}(I_{s_n}). \tag{8.141}$$

Por inducción en n demostraremos que $I_{s_0 s_1 \cdots s_n}$ es distinto del vacío. $I_{s_0} \neq \phi$ puesto que es I_- o I_+. Supongamos que $I_{s_0 s_1 \cdots s_{n-1}} \neq \phi$ y tomemos un punto $x \in I_{s_0 s_1 \cdots s_{n-1}}$, entonces $L_\mu(x) \in I_{s_0 s_1 \cdots s_{n-1}-}$ o $L_\mu(x) \in I_{s_0 s_1 \cdots s_{n-1}+}$, si aquel que resulte coincide con $I_{s_0 s_1 \cdots s_n}$, entonces x es el punto que buscamos, si no coincide entonces $L_\mu^{-1}(L_\mu^{(n)}(x))$ consistirá de dos puntos uno en I_- y el otro en I_+, necesariamente uno de ellos será el punto que buscamos.

Tenemos en adición que

$$I_{s_0 s_1 \cdots s_n} \subset I_{s_0 s_1 \cdots s_{n-1}}, \tag{8.142}$$

por lo que

$$I_\infty = \bigcap_{n=0}^{\infty} I_{s_0 s_1 \cdots s_n} \neq \phi. \tag{8.143}$$

Finalmente $x \in I_\infty$ tendrá el itinerario $s_0 s_1 s_2 \cdots$, lo que demuestra que h es sobre y por tanto que llegamos a la conclusión de que es biyectiva.

Pasemos a demostrar que h es continua. Tomemos una $\epsilon > 0$ arbitraria y tomemos un entero $n > 0$ tal que

$$\frac{1}{2^n} < \epsilon. \tag{8.144}$$

La continuidad del mapeo logístico implica la de sus potencias, por tanto dado un punto $x \in \Lambda$ existe un número $\delta > 0$ tal que si $y \in \Lambda$ y

$$|x - y| < \delta, \tag{8.145}$$

entonces los primeros n términos del itinerario de x y el de y coinciden, por tanto, recordando la métrica de $\sum_2^+$ dada por ((8.131)),

$$d(h(x), h(y)) < \frac{1}{2^n} < \epsilon \tag{8.146}$$

y el mapeo h será continuo. La continuidad de h^{-1} se demuestra de manera similar. Concluimos que $h : \Lambda \to \sum_2^+$ es un homeomorfismo.

En adición h establece una conjugación topológica entre el mapeo logístico restringido a Λ y un corrimiento de Bernoulli en $\sum_2^+$, veamos esto. Tomemos un punto $x \in \Lambda$ con itinerario $s_0 s_1 s_2 \cdots$, esto es $h(x) = s_0 s_1 s_2 \cdots$, consideremos el punto $x' = L_\mu(x)$, su itinerario será $s_1 s_2 \cdots$, esto es $h(x') = h(L_\mu(x)) = s_1 s_2 \cdots$, pero esta última secuencia es un corrimiento de Bernoulli de $s_0 s_1 s_2 \cdots$, o lo que es lo mismo, es $\sigma(h(x))$. Este argumento se puede repetir en dirección inversa por lo que

$$h \circ L_\mu = \sigma \circ h, \tag{8.147}$$

como queriamos demostrar.

Concluimos que, por contener un subsistema topológicamente conjugado a un corrimiento de Bernoulli, el sistema dinámico generado por el mapeo logístico para $\mu > 2 + \sqrt{5}$ es caótico, esto es tiene las dos propidades que caracterizan a estos sistemas dinámicos. Este procedimiento, que únicamente hemos ejemplificado para el caso del mapeo logístico, es típico de la dinámica simbólica y se utiliza frecuentemente para encontrar propiedades de los sistemas dinámicos.

En el caso de los sistemas hamiltonianos debemos ir a mapeos en más de una dimensión, ya que los mapeo de Poincaré para sistemas de dos grados de libertad son del plano en el plano y requerimos al menos dos grados de libertad para que existan sistemas no integrables y por tanto con posibilidad de ser caótico. Pasemos, para poder aplicar algunos resultados a la mecánica, a estudiar un mapeo importante en dos dimensiones, el llamado herradura de Smale o simplemente herradura.

El mapeo de la herradura fue ideado por Smale como un ejemplo de un sistema estructuralmente estable que contiene un número infinito de órbitas periódicas. Esto es de gran importancia para la teoría de los sistemas dinámicos ya que se pensaba que los sistemas de comportamiento 'regular', aquellos que contienen sólo un número finito de órbitas periódicas hiperbólicas y no presentan tangencias homoclínicas, constituían un subconjunto denso dentro del conjunto de todos los sistemas dinámicos diferenciales; la construcción de este ejemplo nos muestra que esto es falso, pues se trata de un sistema no 'regular', esto es caótico, que tiene una vecindad compuesta únicamente de sistemas caóticos, lo que significa que el comportamiento 'regular' no es genérico o equivalentemente que los sistemas caóticos no son una excepción o una curiosidad insignificante, sino que constituyen un amplio subconjunto de todos los sistemas dinámicos. Desde el punto de vista de la mecánica el

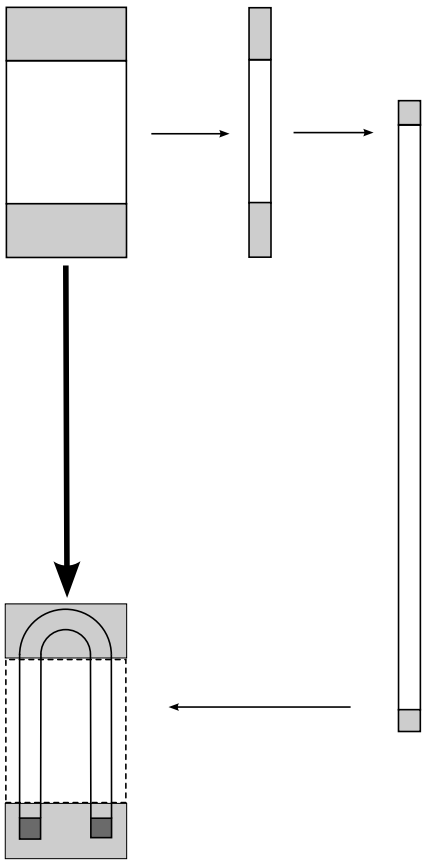


Figura 8.20: El mapeo de Smale o de la herradura. Con flechas delgadas se muestran los pasos que ilustran su construcción. El mapeo completo está indicado por la flecha gruesa.

mapeo o mejor dicho la clase de mapeos de la herradura son importantes ya que con ellos se puede demostrar, como veremos, que los sistemas no integrables son caóticos. Pasemos pues a estudiar estos mapeos.

Tomemos una región de dos dimensiones, podemos pensarla en R^2, y tracemos en ella un rectángulo con otros dos rectángulos trazados sobre lados opuestos del primero, como se muestra en la figura (8.20) en la que hemos dejado en blanco el primer rectaángulo y sombreado los segundos[10]. Ahora pensemos en nuestra región como algo deformable y así describir el mapeo como una deformación de ella en la que cada punto se desplaza; en

[10]Estas figuras constituyen únicamente una guía para poder describir el mapeo.

realidad el desplazamiento de cada punto debemos verlo como la posición a la que el mapeo lo manda. En un primer paso aplicamos una contracción en la dirección horizontal por alguna proporción menor que un medio con esto el trazo guía queda como se muestra en la figura. A continuación aplicamos una expansión en la dirección vertical de tal suerte que el rectángulo blanco quede al menos tres veces más largo, pero aplicamos una contracción en una vecindad de los extremos para que los rectángulos sombredos queden reducidos, ver la figura. Finalmente doblamos de tal manera que se forme una figura de herradura que quede en la posición, con respecto al lugar que ocupaba el trazado original, que se muestra en la figura; el mapeo de herradura es el resultado final de estos tres pasos. Se puede demostrar que es posible hacer estas deformaciones por medio de un difeomorfismo de la región en si misma, esto es con desplazamientos diferenciales e invertibles[11]. Hay desde luego un sinnúmero de realizaciones concretas de mapeos que concuerdan con esta descripción, pero, como veremos, para los propósitos que tenemos todas son buenas, por lo que nos quedamos a este nivel de descripción.

Al aplicar el mapeo nuevamente tendremos la imagen que se muestra en la figura (8.21) y al ir iterando el mapeo se irían generando figuras en las que el rectángulo blanco inicial estaría cruzado cada vez por un número mayor de franjas verticales, de hecho en cada iteración se duplica este número y las franjas se van haciendo cada vez más delgadas. Lo que nos interesa, de manera parecida a lo que nos interesó en el caso ilustrativo del mapeo logístico, es encontrar los puntos del rectángulo blanco que permanecen dentro de éste al iterar el mapeo; vemos que son precisamente los que están en estas franjas y únicamente los que están en estas franjas, ya que, si observamos el rectángulo blanco original, vemos que lo podemos dividir en cinco franjas horizontales, una inferior que se mapea en el pequeño rectángulo que queda dentro del rectángulo sombreado inferior justo debajo de la franja vertical, la siguiente franja horizontal es la que se mapea en la franja vertical de la izquierda, la franja intermedia se mapea en la parte curvada que queda en el rectángulo sombreado superior, la siguiente franja se mapea en la franja vertical derecha y la última se mapea en el pequeño rectángulo que queda justo debajo de la franja vertical derecha. Por otra parte el rectángulo sombreado inferior se mapea en un pequeño rectángulo dentro de él mismo, por lo que al ir iterando se ira contrayendo en un punto, esto es todo punto del rectángulo sombreado inferior o que en alguna de las iteraciones caiga en él, convergerá

[11] En adición a la dimensión aquí encontramos una diferencia importante con el mapeo logístico, este último no es invertible.

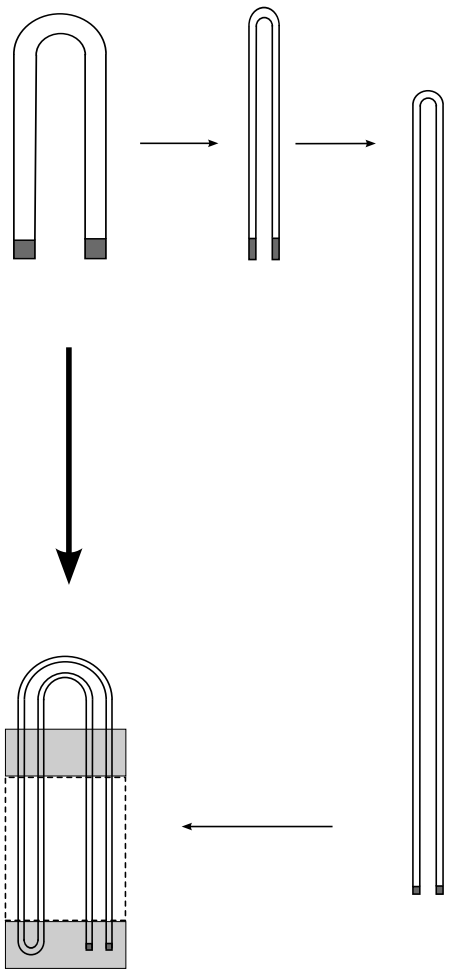


Figura 8.21: Al aplicar dos veces el mapeo de Smale la región blanca original queda cruzada cuatro veces por la región en la que se mapea.

a ese punto fijo del mapeo. El rectángulo sombreado superior se mapea en uno pequeño contenido en el inferior. Vemos, al conjuntar todo esto que, con la excepción de los puntos en las dos franjas horizontales que se mapean en las verticales, todos los demás puntos convergerán, al irse iterando el mapeo, al punto fijo que se encuentra en el rectángulo sombreado inferior.

Concluimos que los únicos puntos del rectángulo blanco original cuya órbita permanece siempre en ese rectángulo forman un conjunto de Cantor formado por lineas verticales, esto es el producto cartesiano de una reta vertical por un Cantor.

Pasemos a estudiar el mapeo inverso o lo que es lo mismo las órbitas en el sentido negativo. Como es claro de la figura (8.22) las dos zonas verticales que se formaron en el rectángulo blanco al hacer la primera iteración, provienen de dos franjas horizontales que intersectan a las verticales en cuatro cuadros, los puntos dentro de estos cuadros son los únicos del rectángulo blanco que se mapean y provienen de ese rectángulo. Si pasamos a iterar el mapeo y su inverso una vez más, las dos franjas verticales se subdividen, como hemos visto, en dos cada una, por su parte podemos ver que la aplicación del mapeo inverso produce en cada franja horizontal dos subfranjas correspondientes a los puntos que se mapean según el mapeo inverso dentro del rectángulo blanco; la intersección subdivide los cuatro cuadros del paso anterior en cuatro cuadrados cada uno, como se ve en la figura; es claro que al ir iterando en ambos sentidos se genera un conjunto de Cantor que es un subconjunto invariante ante el mapeo de la herradura, en este caso ante la iteración del mapeo en ambos sentidos, esto es, está totalmente formado por órbitas completas y no hay órbitas completas totalmente incluidas en el rectángulo blanco que no estén incluidas en ese conjunto de Cantor; denotaremos este conjunto de Cantor invariante con la letra Λ.

De manera análoga a lo que se hizo en el caso del mapeo logístico podemos asociar a cada punto $x \in \Lambda$ un itinerario por medio de asignar a cada una de las dos franjas verticales un símbolo, digamos cero o uno, al iterar el mapeo la órbita de x irá pasando por estas franjas, definimos el itinerario de x como la sucesión de símbolos cero o uno que se genera según la órbita vaya cayendo en una u otra de las franjas. En este caso el mapeo es invertible, por lo que la órbita se extiende en ambas direcciones, podemos por tanto asociar a cada una de la franjas horizontales dos símbolos, por comodidad usaremos los mismos, así obtendremos otra secuencia según la órbita de x en sentido inverso vaya cayendo en una u otra franja horizontal. Obtenemos así dos secuencias, definimos el itinerario de x como la secuencia infinita por la izquierda y por la derecha que se genera al colocar una tras otra las dos secuencias ya obtenidas, colocando la correspondiente a la dirección inversa

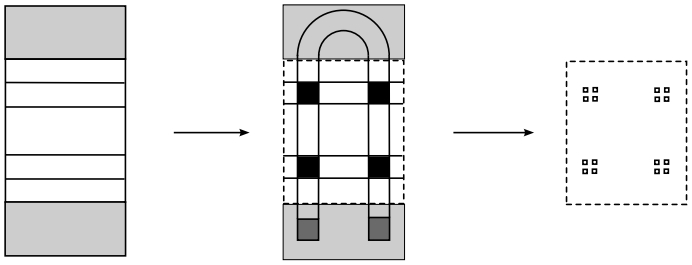


Figura 8.22: Las zonas cuyos puntos se mapean dentro de la región blanca tanto en el sentido inverso como en el directo van formando un conjunto cantor a medida que aumentamos la potencia de la iteración. En el límite obtendremos un conjunto de cantor invariante constituido por todas las órbitas que permanecen en la región blanca.

a la izquierda y la correspondiente a la dirección directa a la derecha de una marca, digamos un punto entre las dos. Un itinerario se verá así

$$\cdots 0010100011.1100010110\cdots . \tag{8.148}$$

Tomemos ahora el espacio $\sum_2$ formado por todas las secuencias biinfinitas de ceros y unos con una marca $S = \cdots s_{-2}s_{-1}.s_0s_1s_2\cdots$ y tomemos una métrica en él definida por

$$d(S,S') = \sum_{i=-\infty}^{\infty} \frac{\delta(s_i, s_i')}{2^{|i|}}, \tag{8.149}$$

donde $\delta(s_i, s_i')$ es cero si $s_i = s_i'$ o uno en caso contrario.

El mapeo $\sigma : \sum_2 \to \sum_2$ definido por

$$\sigma(\cdots s_{-2}s_{-1}.s_0s_1s_2\cdots) = \cdots s_{-2}s_{-1}s_0.s_1s_2\cdots , \tag{8.150}$$

que corre la secuencia un lugar a la izquierda (o lo que es lo mismo corre la marca un lugar a la derecha) será en este caso el corrimiento de Bernoulli. Es importante hacer ver que en este caso el corrimiento de Bernoulli si tiene un inverso que consiste en recorrer la marca un paso hacia la izquierda.

Por un procedimiento totalmente análogo a lo hecho para el caso del mapeo logístico, podemos hacer ver que la correspondencia que los itinerarios establecen entre el conjunto de Cantor Λ y $\sum_2$ es un homeomorfismo y que éste conjuga topológicamente al mapeo de herradura restringido a Λ con el corrimiento de Bernoulli en $\sum_2$. Es también simple hacer ver que σ es caótico, por lo que la herradura sobre Λ también lo será.

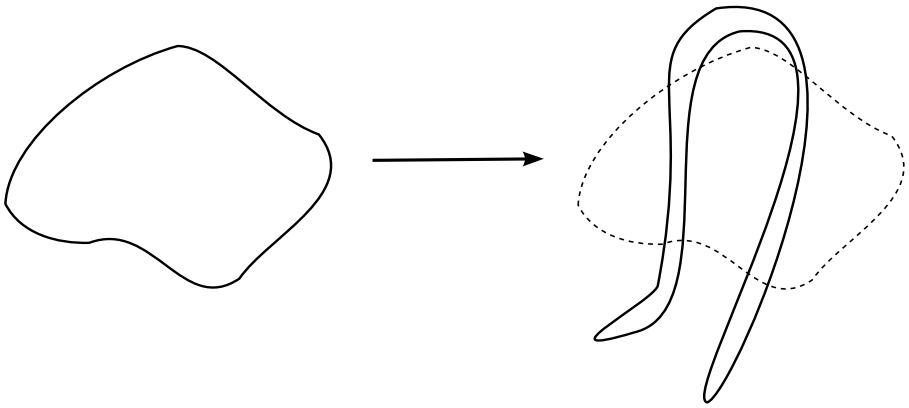


Figura 8.23: Un mapeo que presente la deformación general ilustrada en la figura tendrá el comportamiento dinámico de un mapeo de herradura de Smale.

Hemos así construido una clase de mapeos en dos dimensiones que contienen un subconjunto invariante sobre el cual son topológicamente conjugados a un corrimiento de Bernoulli sobre un espacio de secuencias biinfinitas de dos símbolos. Estos mapeos pueden ser generalizados a mapeos del tipo que se muestra en la figura (8.23) sin que pierdan las propiedades básicas que el de herradura representa.

El mapeo de herradura que utilizamos no preserva área, de hecho sobre la región sombreada inferior es una contracción, esto impide su aplicación a los sistemas hamiltonianos, ya que estos preservan área o, en general, volumen. La razón principal para tomar un caso que no preserva área fue la tener un argumento sencillo y claro para hacer ver que puntos de la región blanca que salen de ella no regresan jamás y que por tanto Λ debe estar totalmente contenido en dicha región. Smale demostró que se puede construir un mapeo de herradura que preserva área, esto es que expande en la misma proporción en que contrae, y que tiene todas las propiedades del que utilizamos en nuestra presentación. Las demostraciones se complican, por lo que, en función de la claridad en la exposición, nos limitamos al caso presentado. También es posible generaliar esto a más dimensiones, pero no lo haremos aquí.

Regresemos ahora a los sistemas hamiltonianos y recordemos su situación. Tenemos intersticios entre los toros compuestos de trayectorias cuasiperiódicas, en estos hay trayectorias periódicas aisladas provenientes de los toros conmensurables rotos que pueden ser estables o inestables. El mapeo de Poincaré o su composición presenta en los intersticios puntos fijos que son o

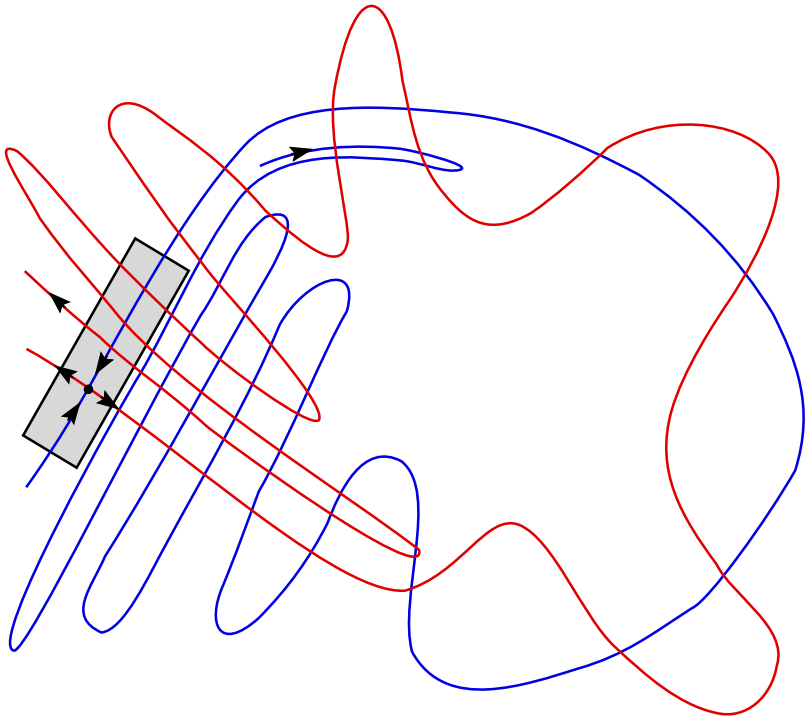


Figura 8.24: Una maraña homoclínica en la que se ha trazado un pequeño rectángulo en torno al punto fijo.

elípticos o hiperbólicos; en torno a los puntos fijos hiperbólicos las variedades estables e inestables al intersectarse homoclínicamente forman las llamadas marañas homoclínicas, de las que representamos una en la figura (8.24).

Observamos que en la figura se ha trazado un pequeño rectángulo que contiene al punto fijo. Al aplicar el mapeo este rectángulo se mapeará en una región que seguirá conteniendo al punto fijo y estará cruzada por un pedazo de la variedad estable y uno de la inestable, naturalmente esta región, cuya área debe ser la del rectángulo original, se alargará a lo largo de la variedad inestable y se comprimirá en la dirección de la estable, como se muestra en la figura (8.25).

Al aplicar nuevamente el mapeo ocurrirá otro tanto de lo mismo, por lo que, al ir iterando, la región se hará cada vez más delgada y se extenderá, cada vez más siguiendo el contorno de la variedad inestable. Vemos que necesariamente después de un número finito de iteraciones se presentará la situación que se muestra en la figura (8.26); la región al seguir a la variedad

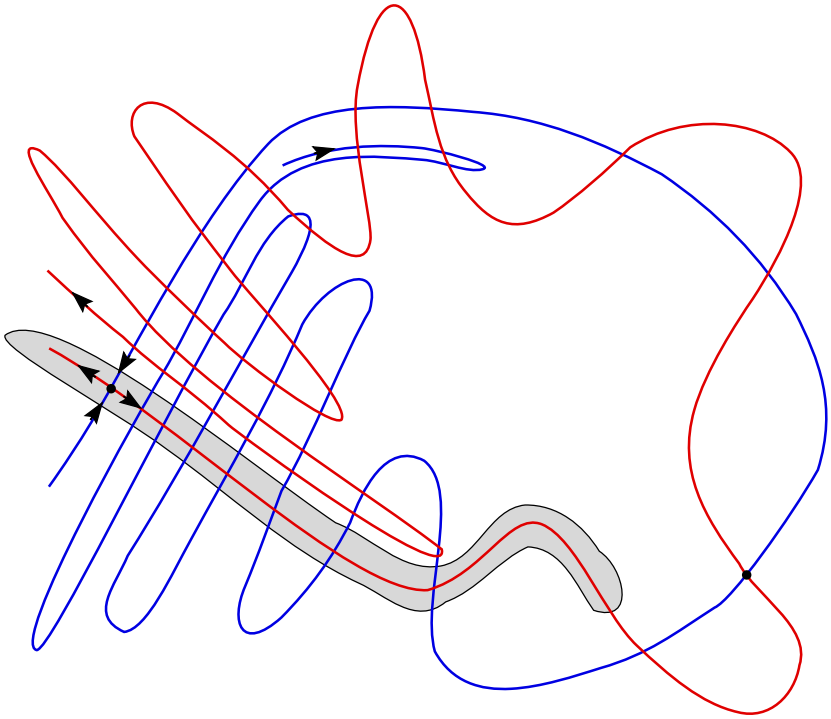


Figura 8.25: Al iterar el mapeo unas cuantas veces el pequeño rectángulo de la figura (8.24) se ha acortado en la dirección de la variedad estable y alargado en la dirección de la variedad inestable. Por razones de la ilustración no se ha acortado la región en la misma proporción en que se ha estirado como señala el teorema de Liuville, en realidad la región debería estar más pegada a la variedad inestable. Por razones de la extensión de esta figura, hemos dejado corta la rama que tiende a estrecharse y recorrer la variedad inestable por el lado izquierdo de la figura, en realidad esta rama también se alargará y se pegará a la variedad.

inestable regresa al entorno del punto fijo e intersecta al rectángulo inicial y, dado que la región continua conteniendo al punto fijo, el mapeo de Poincaré compuesto un número suficiente de veces contendrá una herradura de Smale. Ahora como el mapeo de Poincaré o alguna de sus potencias contienen una herradura de Smale, el sistema dinámico que generan será caótico, por lo que también lo será el sistema hamiltoniano.

En adición a esto podemos ver que no sólo hay una herradura sino que hay un sinnúmero de ellas generadas por la maraña homoclínica. Esto nos muestra la gran complejidad que aparece en los sistemas no integrables, ya que no se tiene un sólo conjunto de Cantor invariable ante el mapeo de Poincaré, se tiene un sinnúmero de ellos. El movimiento correspondiente se verá sumamente irregular, aunque siempre existirán condiciones iniciales para las que el movimiento correspondiente será periódico, pero estas órbitas periódicas estarán aisladas y será muy difícil observarlas ya que forman un subconjunto de medida cero.

8.5. Ejemplos ilustrativos de sistemas no integrables.

Con el objeto de ilustrar el movimiento de los sistemas hamiltonianos no integrables, pasemos a analizar algunos ejemplos. Queda claro que los resultados que obtengamos con esto ejemplos simples se generalizan a otros sistema, en especial a los de mayor número de grados de libertad.

8.5.1. El rotor pateado.

Comenzaremos por el llamado rotor pateado. Consideremos, como sistema integrable, una masa m constreñida a moverse sobre un círculo de radio r. El espacio fase es un cilindro y utilizando coordenadas canónicas (θ, p) la hamiltoniana será

$$H_0 = \frac{p^2}{2mr^2}, \tag{8.151}$$

que corresponde al movimiento circular uniforme con velocidad angular constante $\dot{\theta} = \omega = p/mr^2$. Hagamos actuar una perturbación consistente en la aplicación, cada τ unidades de tiempo, de una fuerza impulsiva sobre la partícula, siempre con la misma intensidad ϵ, en la misma dirección y en el mismo sentido. El efecto de esta fuerza impulsiva es el de provocar un cambio instantáneo en el valor del impulso angular proporcional a la componente tangencial de dicha fuerza. En los intervalos de tiempo en que no actúa la

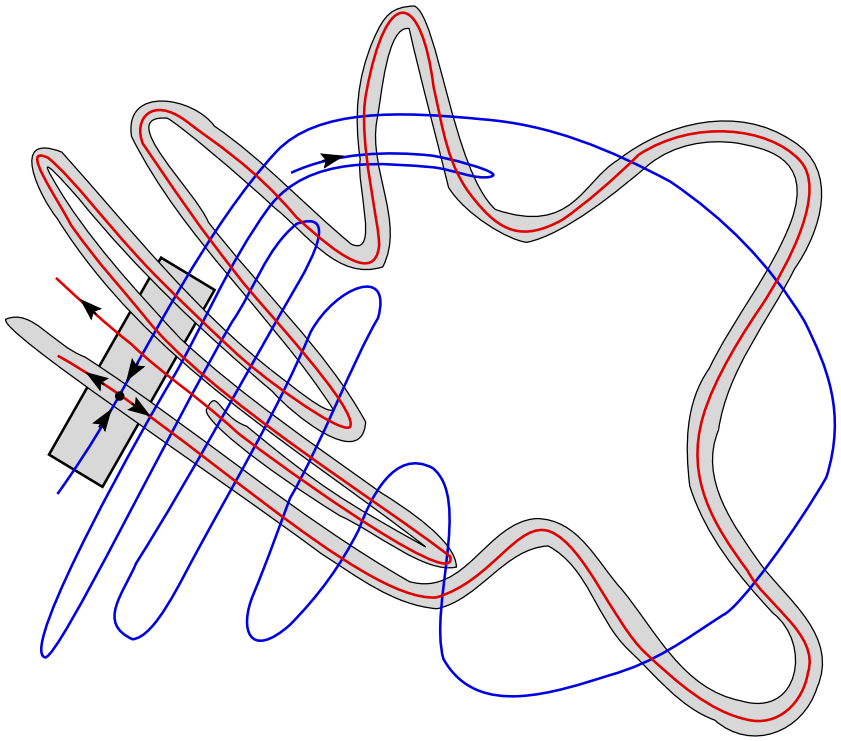


Figura 8.26: Después de un cierto número de iteraciones la región se habrá mapeado en una franja muy estrecha en torno de la variedad inestable e intersectará a la región original dando origen a una herradura de Smale. Al igual que en la figura (8.25) no se ha respetado la proporción entre alargamiento y encogimiento.

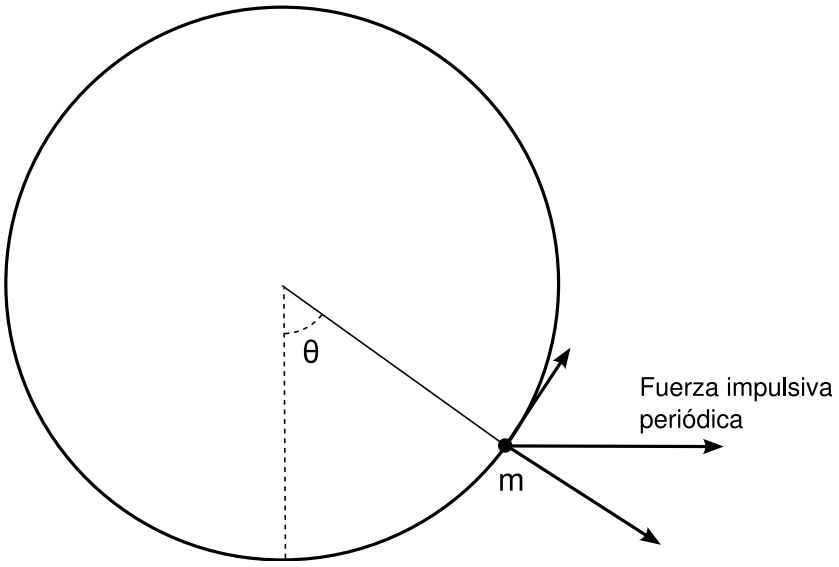


Figura 8.27: El rotor pateado.

fuerza impulsiva, la masa tendrá un movimiento uniforme correspondiente al valor que tenga el impulso.

En el momento en que actúa la fuerza impulsiva la masa estará en una posición $\theta(t)$, por lo que el efecto será el incrementar p por una cantidad

$$\Delta p = \epsilon \cos\theta. \tag{8.152}$$

Esto es, p permanecerá constante hasta que transcurra un período τ, sufrirá un cambio instantáneo Δp y continuará constante hasta que se cumpla un período más. Por su parte θ avanzará conforme a

$$\theta = \theta_0 + \omega(t)t \quad (mod\ 2\pi), \tag{8.153}$$

donde $\omega(t)$ será constante excepto al ocurrir un cambio en el impulso cuando cambiará instantáneamente por la cantidad $\Delta\omega = \Delta p/mr^2$.

Para estudiar este movimiento podemos utilizar una técnica muy parecida a la del mapeo de Poincaré. Como tenemos una perturbación periódica en el tiempo con período τ, podemos tomar un punto en el espacio fase y mapearlo a aquel en que la trayectoria se encontrará τ segundos después. Este mapeo genera un sistema dinámico discreto en el espacio fase cuyas órbitas reflejarán las propiedades de las trayectorias del sistema mecánico. Podemos ver esto como un mapeo de Poincaré si consideramos al tiempo como una variable

circular en el espacio fase extendido de tres dimensiones y el corte transversal lo hacemos perpendicularmente a está variable temporal circular. Veamos cual es este mapeo.

Supongamos que un instante antes de ocurrir la iésima perturbación los valores de las coordenadas son θ_i y p_i, después de transcurrido un período los valores serán θ_{i+1} y p_{i+1} y estarán dados por

$$\theta_{i+1} = \theta_i + \omega_{i+1}\tau = \theta_i + \frac{p_{i+1}}{mr^2}\tau \pmod{2\pi}, \tag{8.154}$$

$$p_{i+1} = p_i + \epsilon \cos\theta_i, \tag{8.155}$$

la última ecuación proviene del hecho de que en el instante en que ocurre la fuerza impulsiva, θ vale θ_i. Tomando, por simplicidad, a la masa, el radio y el período como unitarios, tendremos que el mapeo será

$$\theta_{i+1} = \theta_i + p_{i+1} \pmod{2\pi}, \tag{8.156}$$

$$p_{i+1} = p_i + \epsilon \cos\theta_i. \tag{8.157}$$

Este mapeo se conoce como el mapeo standard y sus órbitas en el espacio fase reflejan las propiedades del rotor pateado.

Cuando ϵ es cero tendremos un movimiento circular uniforme, las trayectorias en el espacio fase serán simplemente los círculos de p constante que constituyen una foliación completa del cilindro por toros de una dimensión. El mapeo ((8.157)) se reduce a

$$\theta_{i+1} = \theta_i + p_{i+1} \pmod{2\pi}, \tag{8.158}$$

$$p_{i+1} = p_i, \tag{8.159}$$

cuyas órbitas serán

$$\theta_n = \theta_0 + np_0 \pmod{2\pi}, \tag{8.160}$$

$$p_n = p_0. \tag{8.161}$$

Una órbita será una serie de puntos en el cilindro colocados sobre el círculo correspondiente al valor p_0 del impulso; si p_0 es una cantidad conmensurable con 2π la órbita será periódica y consistirá en un número finito de puntos; si, por el contrario, es inconmensurable, la órbita será cuasiperiódica y contendrá una infinitud de puntos que integrarán un subconjunto denso del círculo correspondiente. Es interesante observar que el mapeo standard para $\epsilon = 0$ representa toros conmensurables e inconmensurables, mientras que sabemos que en el sistema continuo todos son periódicos ya que en un grado de libertad no puede haber soluciones cuasiperiódicas. Ocurre que en la construcción

del mapeo introdujimos un período τ que se convierte en nuestra referencia y la conmensurabilidad es relativa a este período. Cuando ϵ es cero este período es artificial, pero es esencial en el momento en que ϵ deja de ser cero; como la dinámica está determinada por las resonancias con este período τ el mapeo standard las refleja y, como veremos, los toros conmensurables se rompen al dejar ϵ de ser cero.

El mapeo standard con $\epsilon \neq 0$ tiene puntos fijos, los que podemos obtener a partir de ((8.157)) por medio de

$$\theta_{i+1} = \theta_i + p_i + \epsilon \cos(\theta_i) = \theta_i \pmod{2\pi} \tag{8.162}$$

y

$$p_{i+1} = p_i + \epsilon \cos(\theta_i) = p_i. \tag{8.163}$$

Estas ecuaciones tienen como soluciones $\theta = \pi/2, 3\pi/2$ y $p = 0$ o un múltiplo entero de 2π. Las dos soluciones con $p = 0$ corresponden a tener el rotor parado en los dos puntos en los que la fuerza impulsiva es radial y por tanto no tiene efecto, los demás son aquellos en que el rotor tiene una velocidad angular tal que hace que la masa pase exactamente por los puntos $\theta = \pi/2$ o $\theta = 3\pi/2$ en los instantes en que se aplica la fuerza impulsiva, la que, al ser radial en esos puntos, no tiene efecto.

Un ejemplo de las órbitas periódicas que existen independientemente del valor de ϵ lo tenemos cuando $p = \pi$ y θ pasa en cada iteración de valer $\pi/2$ a valer $3\pi/2$. Todas las órbitas que como ésta son periódicas debido a que el valor de p hace que se sincronice la ocurrencia de la fuerza impulsiva con alguna de las posiciones $\theta = \pi/2, 3\pi/2$, provienen de la ruptura de toros conmensurables con el período τ del sistema integrable. Es fácil ver que todas estas órbitas periódicas resultan inestables.

En las figuras (8.28) a (8.32) el espacio fase cilíndrico se representa sobre un plano; en el eje horizontal está la coordenada angular θ y sobre el vertical el impulso p. Se trazan las órbitas iterando dos mil veces el mapeo standard a partir de las condiciones iniciales. En un mismo diagrama se trazan cien órbitas distintas generadas a partir de condiciones iniciales tomadas al azar.

En la primera figura (8.28) tenemos el caso integrable, esto es $\epsilon = 0$. Podemos observar los toros de la foliación que naturalmente quedan trazados como rectas horizontales. Para algunos se ve la recta total o casi totalmente llena y para otros vemos únicamente un número finito de puntos; los primeros corresponden a toros inconmensurables con la frecuencia, en este caso artificial, con que se aplica la fuerza impulsiva; los segundos corresponden a los conmensurables o, mejor dicho, aquellos que son 'casi conmensurables', pero que, debido al número limitado de iteraciones, no alcanzamos a distinguir

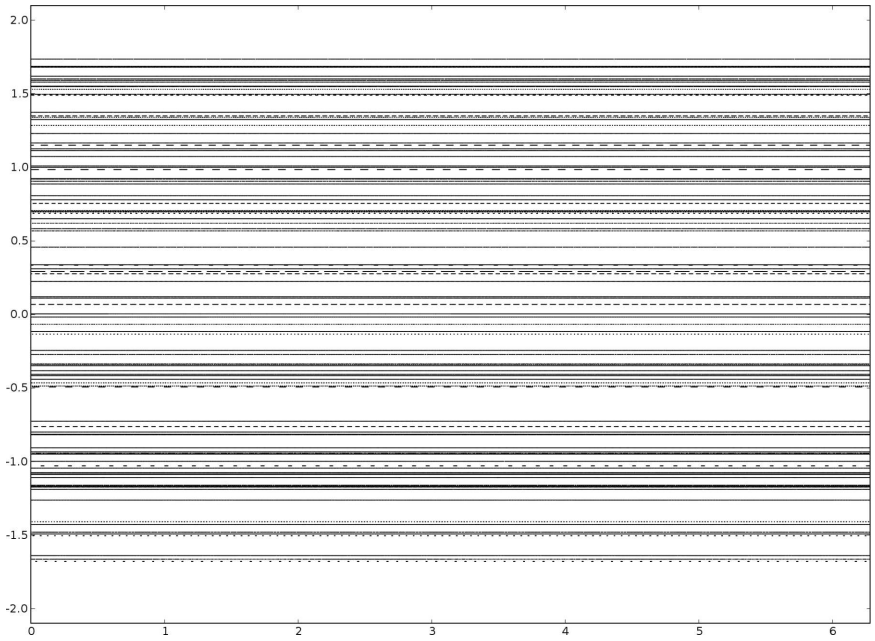


Figura 8.28: El espacio fase del rotor pateado para el caso integrable $\epsilon = 0$. En la figura tenemos cien órbitas cuyas condiciones iniciales se han escogido al azar. El espacio fase es un cilindro, aquí lo hemos cortado para presentarlo en un plano, por lo que hay que identificar el margen izquierdo con el derecho. En las abscisas está la coordenada θ y en las ordenadas el impulso p. Las figuras correspondientes a valores distintos de cero del parámetro ϵ se han elaborado y representado de esta misma manera.

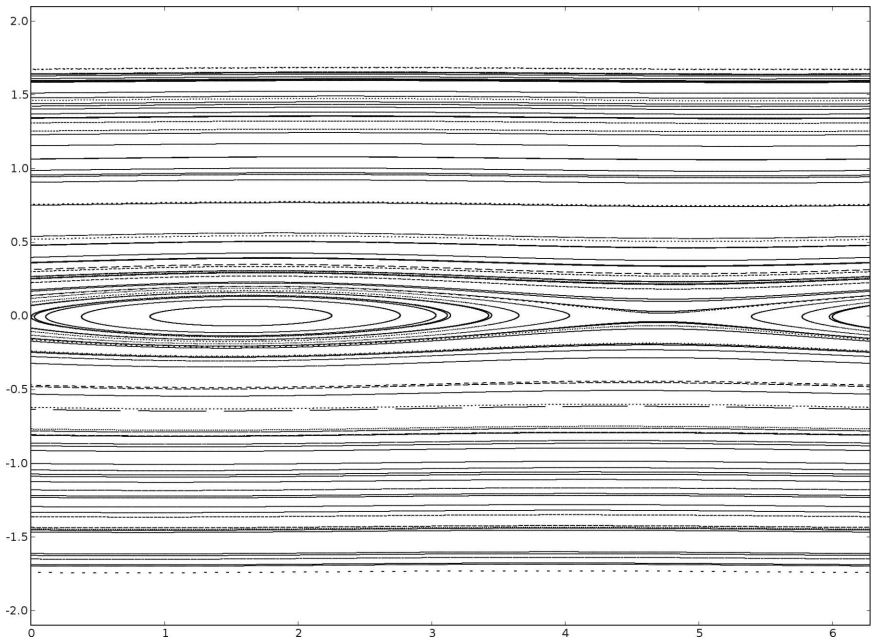


Figura 8.29: El espacio fase del rotor pateado para $\epsilon = 0.01$.

de los realmente conmensurables que nunca se pueden tener con condiciones iniciales aleatorias ya que cubren una medida cero.[12]

En la segunda figura (8.29) el parametro ϵ tiene el valor 0.01, podemos ver la formación de toros en torno del punto fijo estable o elíptico $(\pi/2, 0)$ y la preservación de muchos de los toros originales que únicamente se deforman y comienzan a delinear el entorno del punto fijo inestable o hiperbólico $(3\pi/2, 0)$.

En la siguiente figura (8.30) con $\epsilon = 0.4$ podemos apreciar ya algunas de las órbitas periódicas por medio de la estructura de toros en su entorno[13]. Observamos ya con claridad el carácter hiperbólico de las órbitas cercanas al punto fijo inestable, pero no se alcanza a ver la complejidad de las órbitas en los entornos de esos puntos.

[12] Hay que tomar en cuenta que se trata de soluciones numéricas necesariamente aproximadas por lo que hay que tomar los términos conmensurable e inconmensurable con la cautela debida.

[13] Naturalmente al ser tomadas las condiciones iniciales al azar no obtendremos ninguna órbita periódica. Podemos detectar la presencia de las estables observando que algunas órbitas trazan las curvas cerradas correspondientes a los pequeños toros que se forman en su vecindad.

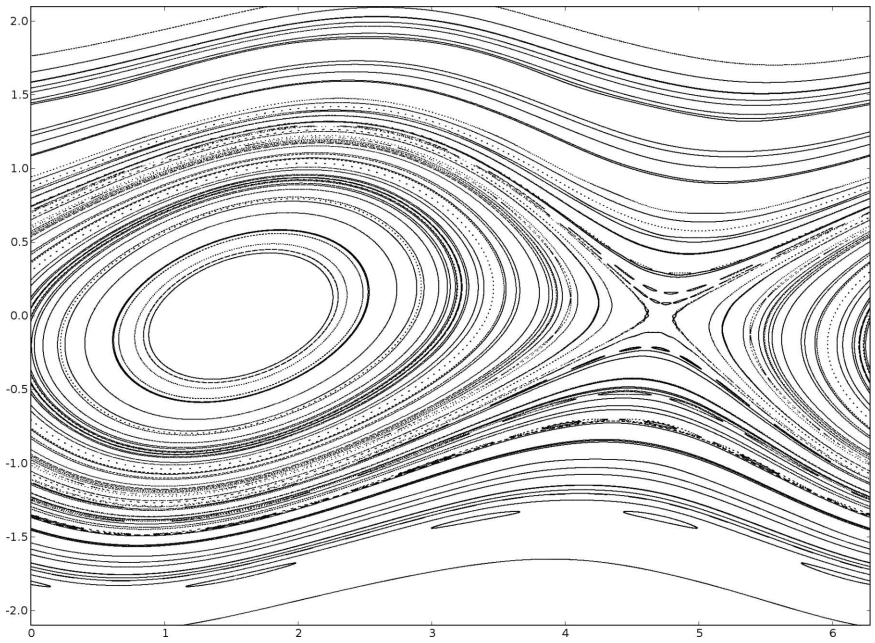


Figura 8.30: El espacio fase del rotor pateado para $\epsilon = 0.4$.

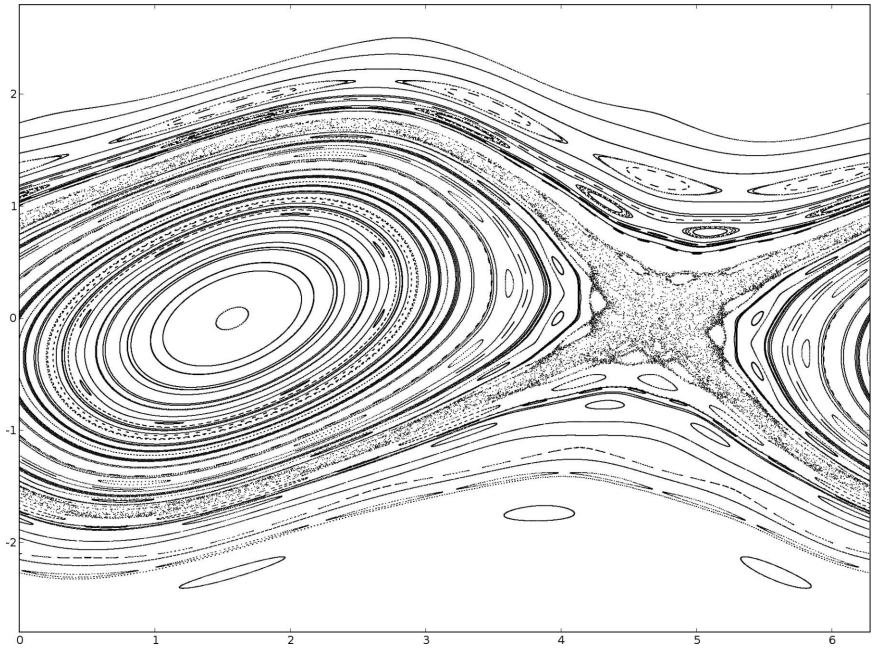


Figura 8.31: El espacio fase del rotor pateado para $\epsilon = 0.7$.

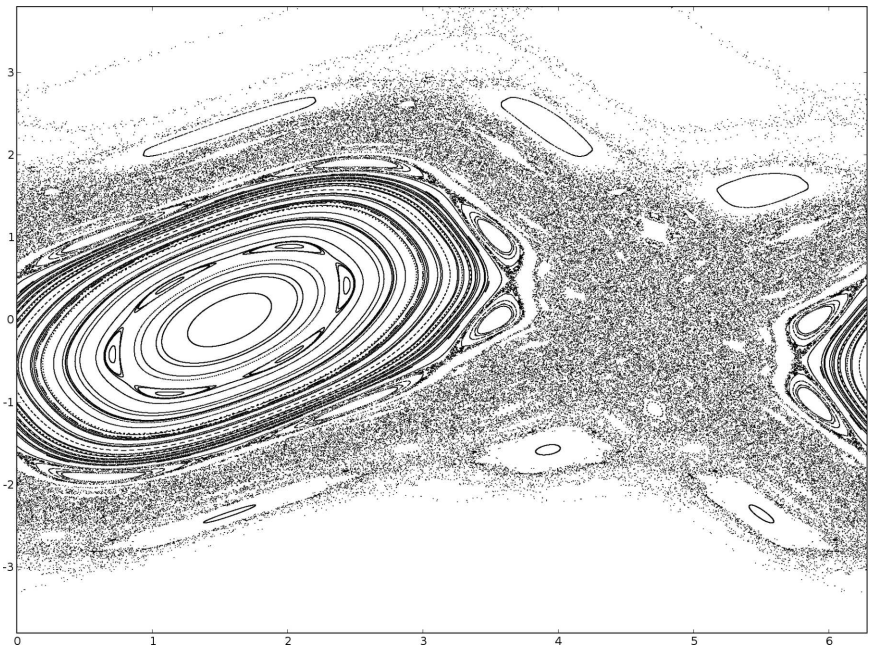


Figura 8.32: El espacio fase del rotor pateado para $\epsilon = 1.1$.

En la figura siguiente (8.31), $\epsilon = 0.7$, ya se revelan plenamente las propiedades del sistema no integrable. Hay órbitas periódicas, tanto estables o elípticas, como inestables o hiperbólicas; vemos la formación de toros en torno a las elípticas; vemos lo complicado que son las órbitas en los entornos de los puntos hiperbólicos, donde predominan las marañas homoclínicas; vemos también como una buena parte de estas órbitas se ve sumamente dispersa y con apariencia aleatoria. Esta figura nos muestra ya con claridad que el rotor pateado es caótico.

Finalmente en la figura (8.32) con $\epsilon = 1.1$ vemos el predominio de las órbitas caóticas al haberse roto la mayoría de los toros 'inconmensurables' del sistema integrable. Notamos que este caso ya nos da una indicación de ergodicidad y mixing.

En este caso es fácil demostrar que el sistema es caótico. El mapeo standard está dado por

$$T: \quad \begin{cases} \theta' = \theta + p + \epsilon \cos\theta \\ p' = p + \epsilon \cos\theta \end{cases} , \tag{8.164}$$

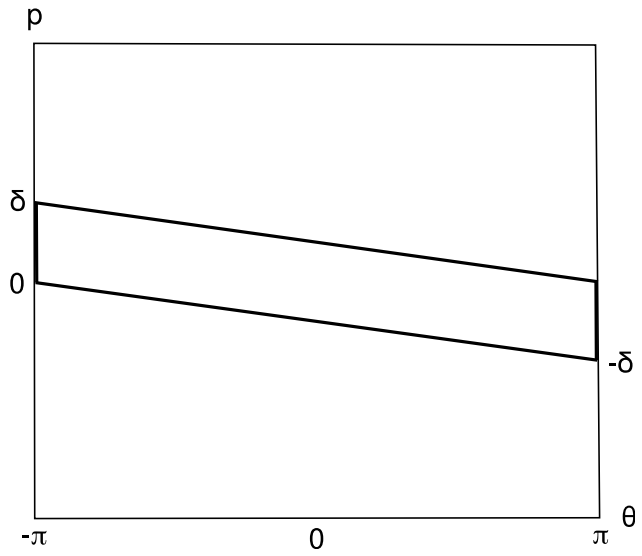


Figura 8.33: Región en el espacio fase del rotor pateado construida según se explica en el texto.

que puede ser visto como la composición de dos mapeos $T = T_2 \circ T_1$ con

$$T_1 : \quad \begin{cases} \theta' = \theta \\ p' = p + \epsilon \cos\theta \end{cases} \tag{8.165}$$

y

$$T_2 : \quad \begin{cases} \theta' = \theta + p \\ p' = p \end{cases} . \tag{8.166}$$

Dado un valor del parámetro ϵ tomemos una cantidad δ tal que $0 < \delta < \epsilon$, tomamos para ilustrar el argumento $\delta = \epsilon/2$, y construimos en el espacio fase cilíndrico la región limitada por el paralelogramo cuyos vértices son $(-\pi, 0)$, $(-\pi, \delta)$, $(\pi, 0)$ y $(\pi, -\delta)$, como se muestra en la figura (8.33).

Al aplicar el mapeo T_1 las rectas verticales permanecen como rectas verticales, mientras que la recta superior dada por

$$p = -\frac{\delta}{2\pi}\theta + \frac{\delta}{2}, \tag{8.167}$$

se transforma en la curva

$$p' = -\frac{\delta}{2\pi}\theta' + \frac{\delta}{2} + \epsilon \cos\theta'; \tag{8.168}$$

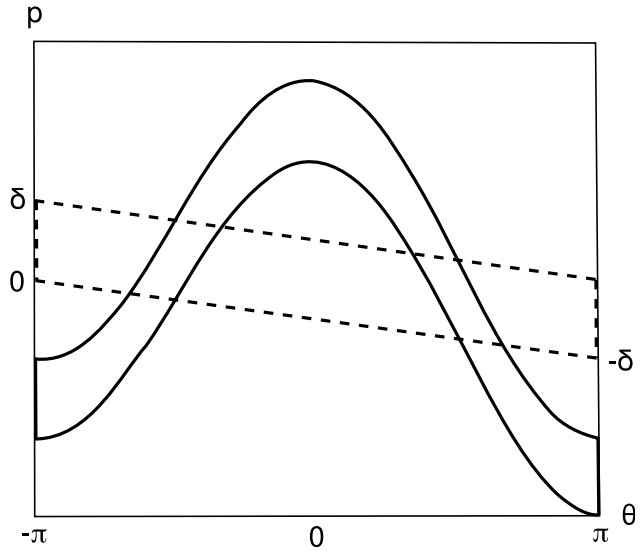


Figura 8.34: Después de aplicar el primer paso de los dos en que se ha dividido el mapeo standard, la región destacada con la linea punteada se transforma en la señalada con la linea continua.

y la inferior dada por

$$p = -\frac{\delta}{2\pi}\theta - \frac{\delta}{2}, \tag{8.169}$$

se transforma en la curva

$$p' = -\frac{\delta}{2\pi}\theta' - \frac{\delta}{2} + \epsilon\cos\theta'. \tag{8.170}$$

La región se transforma entonces en la región que se ilustra en la figura (8.34).

La transformación T_2 es un movimiento de cizalla que deja invariante a la recta $p = 0$, desplaza hacia la derecha la parte superior y hacia la izquierda a la inferior, por lo que al aplicarla a la región de la figura para encontrar el efecto del mapeo standard, tendremos la situación ilustrada en la figura (8.35), en donde hemos incluido el trazo de la región original y tomado en consideración que estamos sobre un cilindro.

Como podemos ver se ha formado una herradura de Smale y por tanto tendremos caos; habrá un subconjunto cantoriano invariante sobre el cual el mapeo standard es topológicamente conjugado a un corrimiento de Bernoulli en el espacio de todas las secuencias biinfinitas de dos símbolos.

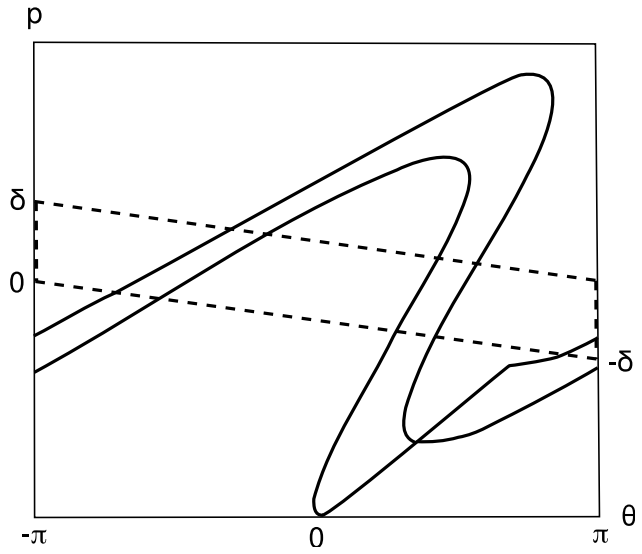


Figura 8.35: Al aplicar el mapeo standard, la región indicada con la linea punteada se mapea en la delineada con la linea continua. Podemos o observar la formación de una herradura de Smale.

Terminamos este análisis del rotor pateado mostrando un método para trazar el enredo homoclínico. Las figuras (8.36) a (8.39) muestran la presencia de las variedades estable e inestable en una vecindad de la intersección homoclínica. Para construirlas se toma una pequeña región (en este caso cuadrada) centrada en el punto fijo hiperbólico y se mapea un cierto número de veces. Se obtiene una región de la misma área pero que se ha encogido en dirección de la variedad estable y se ha estirado a lo largo de la variedad inestable, al hacer esto se perfila la variedad inestable. Vemos como en las primeras iteraciones se comienza a trazar la variedad inestable, a continuación se da la vuelta y la parte que más ha avanzado regresa a las cercanías del punto fijo, finalmente se comienzan a vislumbrar las fuertes oscilaciones de la variedad inestable producidas por la presencia de la intersección homoclínica, podemos observar también la convergencia asintótica al punto fijo. La variedad estable no se visualiza con este procedimiento pero la podemos trazar con un proceso similar pero aplicando el mapeo inverso; al superponer ambas construcciones podemos visualizar la maraña homoclínica como se ve en la última figura.

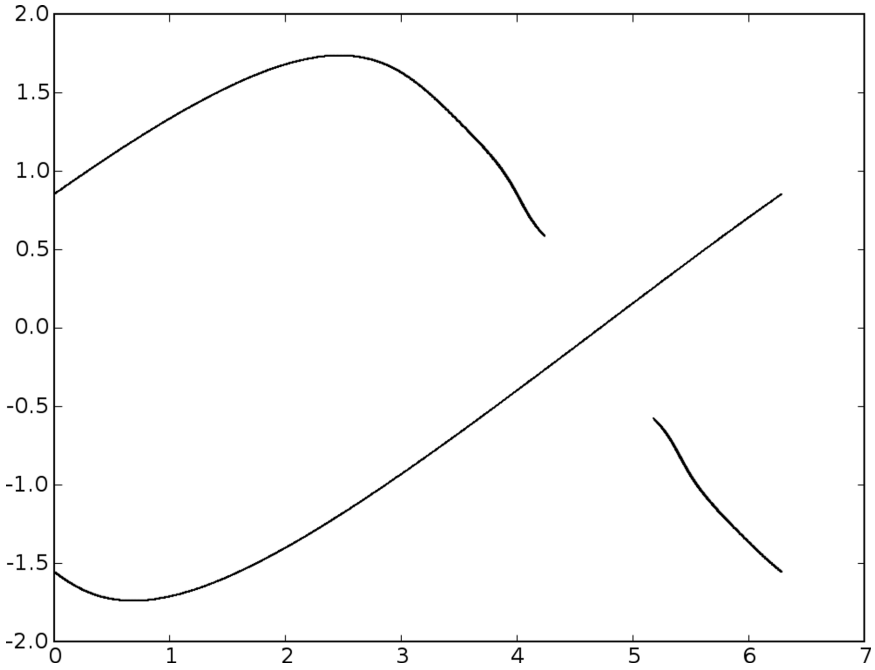


Figura 8.36: Trazado de la variedad inestable del punto homoclínico del mapeo standard.

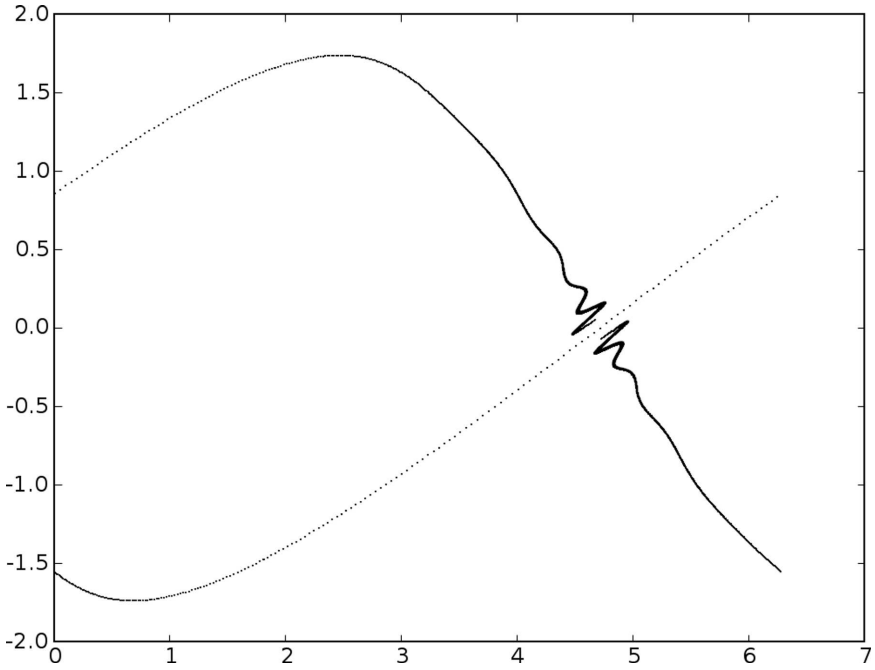


Figura 8.37: Trazado de la variedad inestable del punto homoclínico del mapeo standard.

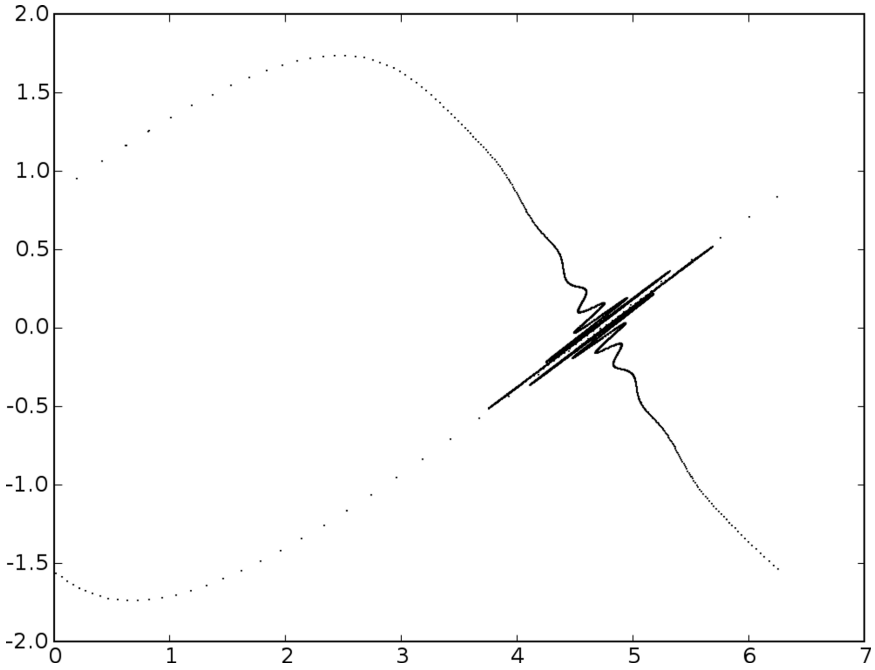


Figura 8.38: Trazado de la variedad inestable del punto homoclínico del mapeo standard.

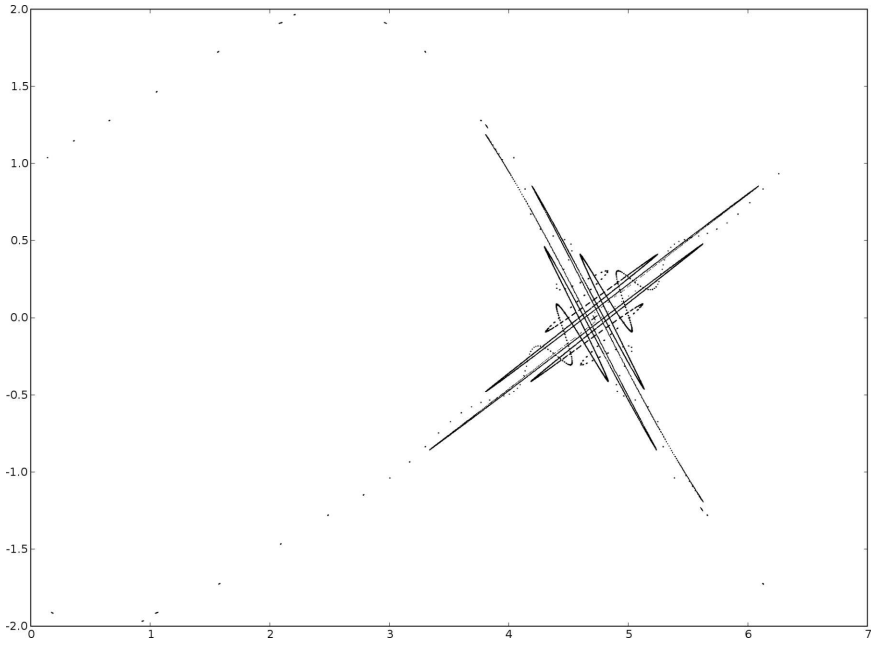


Figura 8.39: La maraña homoclínica del punto fijo inestable del mapeo standard.

8.5.2. La hamiltoniana de Henon y Heiles.

Pasemos a ver un caso continuo, la hamiltoniana de Henon y Heiles. A principios de los años 1960 se había podido medir un buen número de los movimientos de las estrellas cercanas con respecto a la Galaxia. Estos movimientos se dan en un campo gravitacional que, con buena aproximación, podemos considerar como el promedio de la interacción de una estrella con el resto de las que conforman la Galaxia. Ese campo promedio debe tener simetría axial, dada la forma de la Galaxia, entonces en coordenadas cilíndricas centradas en el centro de la Galaxia el potencial debe ser independiente de la coordenada ϕ y la hamiltoniana para una estrella es de la forma

$$H = \frac{1}{2m}(p_r^2 + p_z^2 + \frac{1}{r^2}p_\phi^2) + V(r,z). \tag{8.171}$$

Es inmediato ver que tenemos dos integrales de movimiento, la energía total y la proyección del impulso angular sobre el eje z, p_ϕ, lo que limita el movimiento a una región del espacio fase de cuatro dimensiones. Ahora de no existir otra integral, lo que es muy probable dada la demostración de Poincaré sobre la rareza de los sistemas integrables, y si el movimiento fuese ergódico en esa región, como se pensaba en esos años sobre el movimiento de los sistemas no integrables, las desviaciones standard en las distribuciones del movimiento en r y en z de las estrellas cercanas deberían ser similares; sin embargo se observa una diferencia por un factor dos, difícilmente imputable a errores de medición. Esto constituye un interesante problema, como todos los que muestran una incongruencia entre lo observado y lo que en principio se debiera observar, o bien el potencial no tenía simetría cilíndrica o debe existir otra integral. Ambas cosas parecen poco probables y no se veían razones para pensar en desviaciones de la ergodicidad.

Henon y Heiles, con el objetivo de estudiar el movimiento en un caso atacable numéricamente, construyeron un modelo de dos grados de libertad para la hamiltoniana de una estrella inmersa en el campo gravitacional de una galaxia con simetría en torno a su eje.

$$H = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}(x^2 + y^2) + x^2 y - \frac{1}{3}y^3; \tag{8.172}$$

por comodidad en esta hamiltoniana se han reducido las constantes utilizando unidades en las que resultan unitarias y utilizamos las variables x y y para recalcar que se trata de un modelo y no de un verdadero potencial efectivo para la Galaxia. Este modelo tiene a la energía total como integral de movimiento y es difícil pensar en que tenga otra. En la figura (8.40) se

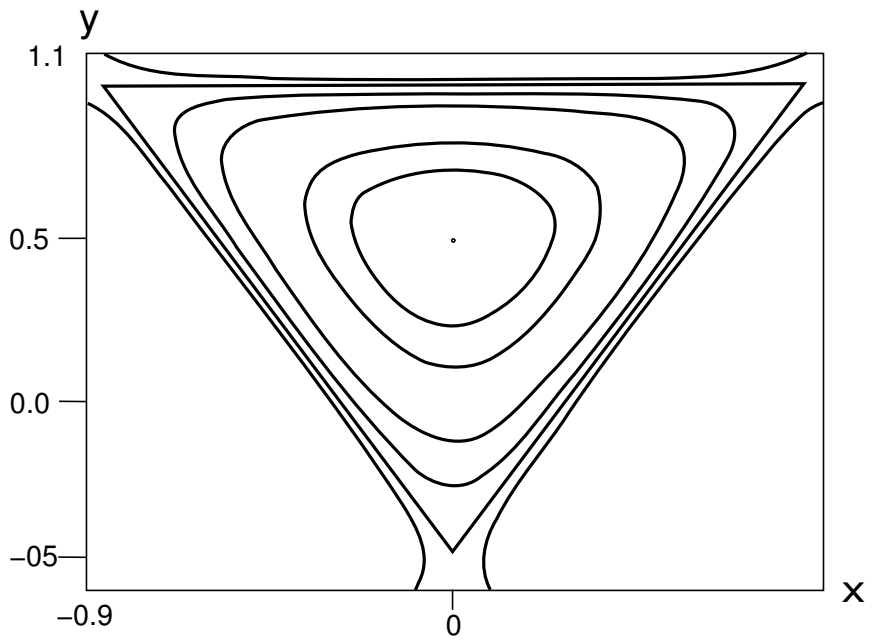


Figura 8.40: Curvas de nivel del potencial de Henon Heiles.

muestran las curvas de nivel del potencial de Henon Heiles, se ve claramente la forma de una cuenca. Esta cuenca está limitada por el triángulo formado por los puntos $(0,1)$, $(\sqrt{3}/2,-1/2)$ y $(-\sqrt{3}/2,-1/2)$; sobre este parteaguas triangular el potencial vale $1/6$. Nos concentraremos en el comportamiento dinámico de este sistema para energías totales en las que el movimiento queda acotado por la cuenca, esto es para $E<1/6$.

En las gráficas se muestran algunas órbitas del mapeo F obtenido de la siguiente manera: se fija un valor de $E<1/6$ y se da el valor inicial cero a la variable x; se toma un punto inicial sobre el plano (y,p_y) que sea compatible con la energía E para $x=0$ y se sigue la trayectoria cuyas condiciones iniciales quedan establecidas por E, y, p_y y $x=0$; p_x queda determinada por estas condiciones; cuando esta trayectoria vuelva a pasar por $x=0$ estará en un punto (y',p_y') que, por definición establece el mapeo F del plano (y,p_y) en si mismo. En todos los casos ilustrados las trayectorias, como es natural, están limitadas por la curva de nivel del potencial correspondiente a la energía total dada; es por tanto importante observar las diferentes escalas que se utilizan en estas gráficas.

En la gráfica (8.41) la energía es 0.01 podemos ver, hasta donde alcanza la resolución, que las órbitas trazan curvas suaves, algunas de ellas en rea-

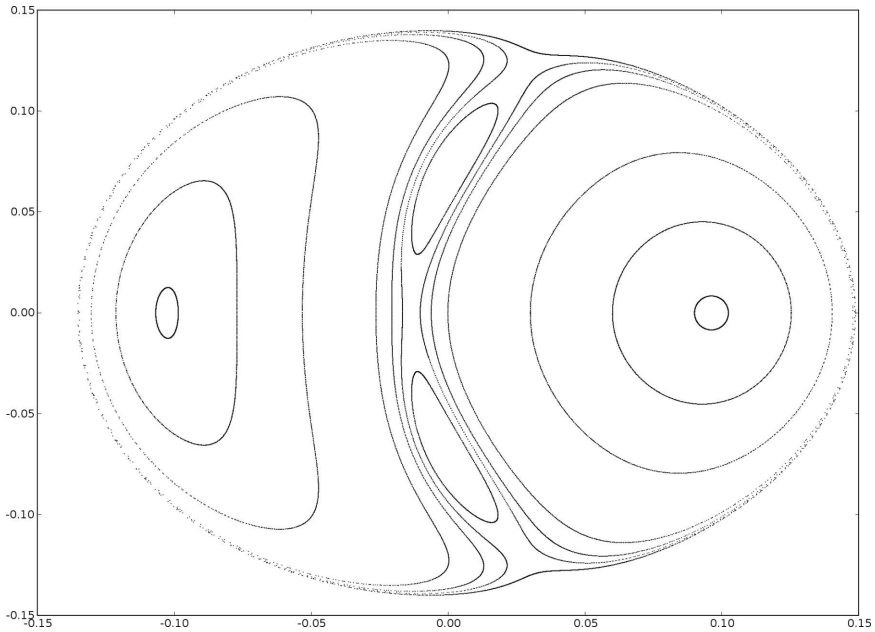


Figura 8.41: El mapeo $F : (y, p_y) \mapsto (y', p'_y)$ del sitema de Henon y Heiles, obtenido como se explica en el texto, para energía total 0.01. Se han graficado ocho trayectorias correspondientes a las condiciones iniciales $x = 0$, $(y, p_y) = (0.0, -0.09)$, $(0.02, -0.12)$,$(0.09, 0.001)$, $(0.06, 0.001)$, $(0.03, 0.001)$, $(0.03, -0.13)$, $(0.01, 0.05)$ y $(0.01, 0.075)$. Estas mismas condiciones iniciales se utilizan para los demás casos ilustrados.

lidad no lo deben ser, pero son muy cercanas de serlo; estas curvas son un reflejo de la presencia en el espacio fase de toros compuestos de haces de trayectorias. Podemos vislumbrar la presencia de puntos fijos inestables y de las separatrices correspondientes. Esto se ve como un sistema muy similar a uno integrable.

En la figura (8.42) la energía es 0.03 podemos ver que comienzan a desdibujarse los toros. En la figura (8.43) para un valor de la energía de 0.1 se han desfigurado los toros aun más. El caso con energía 0.12 se muestra en la gráfica (8.44), notamos claramente la presencia de órbitas que definitivamente no trazan curvas; órbitas que más bien se distribuyen sobre regiones bidimensionales en un proceso que, aunque sólo se han trazado un número finito de puntos, podemos decir que parece ergódico y mixing. Vemos aquí ya el carácter caótico del sistema, pero con la presencia de órbitas que trazan o permanecen muy cercanas a curvas suaves.

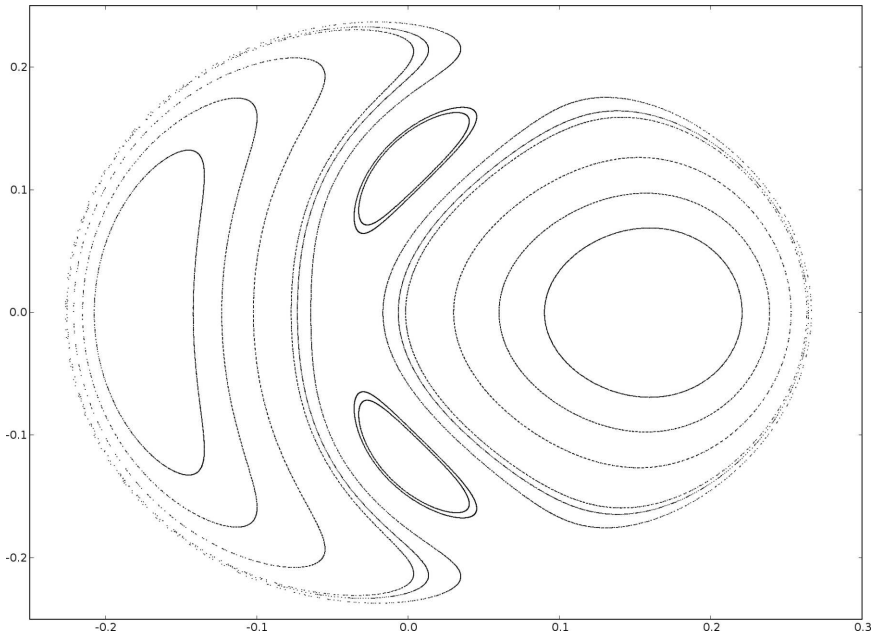


Figura 8.42: El sistema de Henon y Hailes, caso con energía total 0.03.

En la figura (8.45) la energía es 0.15, podemos ver claramente el predominio de las órbitas que no trazan curvas y por tanto el pleno desarrollo de las propiedades caóticas del sistema. Finalmente en la figura (8.46) tenemos el caso con energía 0.166, este caso está muy cercano al límite de 1/6 para el cual las trayectorias dejan de ser acotadas, podemos ver el predominio de las trayectorias caóticas.

Este modelo explica los resultados de las observaciones del movimiento de las estrellas cercanas haciendo ver que, para energías pequeñas, el sistema se comporta como un sistema muy cercano a uno integrable y, por tanto, una mayoría de las trayectorias son cuasiperiódicas y no ergódicas como se suponía anteriormente al teorema de KAM. Este modelo tuvo una gran repercusión en el desarrollo de la teoría de los sistemas dinámicos, pues mostraba, con un ejemplo numéricamente accesible, las características básicas de los sistemas no integrables poco entendidas en los tiempos en que se estudio este potencial por Henon y Heiles.

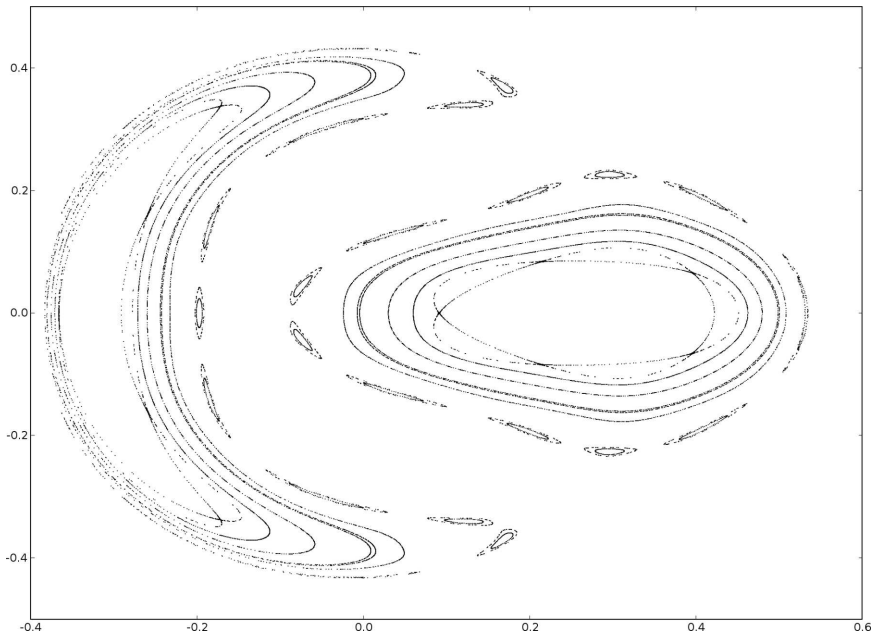


Figura 8.43: El sistema de Henon y Hailes, caso con energía total 0.1.

8.5.3. Billares.

Una clase de sistemas muy estudiada es la que se integra con los billares. Se trata del problema general de una partícula libre encerrada en una región del espacio. La partícula rebota elásticamente al chocar con la frontera de dicha región.

En el caso de que la región sea circular tendremos que, debido a la simetría ante rotaciones por un eje que pasa por el centro de la región, el impulso angular es una integral de movimiento, la que aunada a la energía total nos hace ver que el sistema es integrable. En adición a esto el espacio fase es acotado, por lo que los movimientos son periódicos o cuasiperiódicos y existen variables de acción y ángulo.

En coordenadas polares la hamiltoniana es

$$H = \frac{1}{2m}\left(p_r^2 + \frac{1}{r^2}p_\theta^2\right) = E. \tag{8.173}$$

Las integrales de movimiento son E y p_θ. Las variables de acción serán

$$J_\theta = \frac{1}{2\pi}\oint p_\theta d\theta \tag{8.174}$$

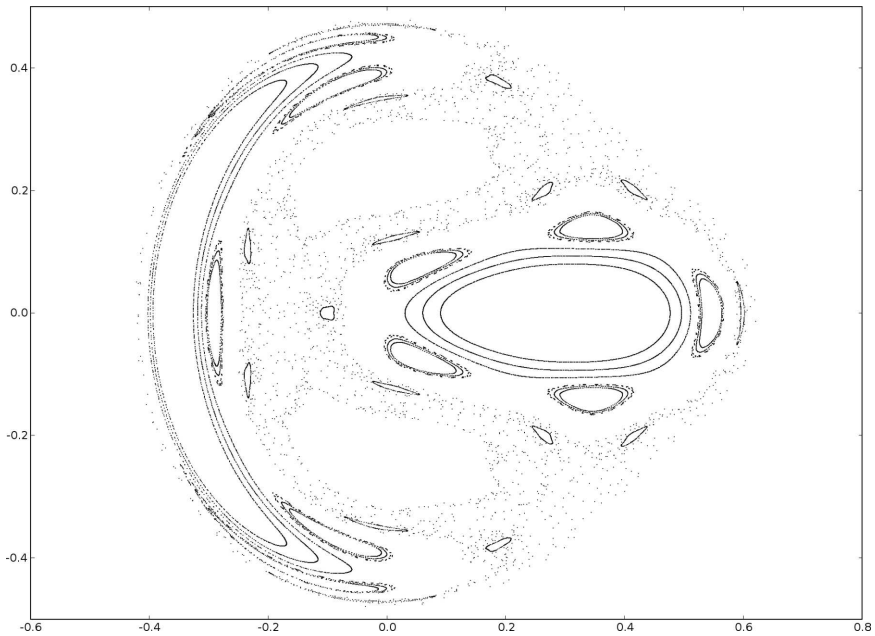


Figura 8.44: El sistema de Henon y Hailes, caso con energía total 0.12.

y

$$J_r = \frac{1}{2\pi} \oint p_r dr. \tag{8.175}$$

Como $p_\theta = mr^2\dot{\theta}$ es constante, $\dot{\theta}$ será monótona y

$$J_\theta = p_\theta. \tag{8.176}$$

Para calcular J_r notamos que la coordenada r recorre un ciclo al pasar del valor R al $R_{\text{mín}}$ y al R de vuelta, en el tiempo en que p_r pasa de p_R, el valor de p_r cuando $r = R$, a 0 y de nuevo a p_R, siendo p_r negativo en el primer tramo y positivo en el segundo, pero con valores simétricos. Por ello la integral cíclica será el doble de la integral de $R_{\text{mín}}$ a R, esto es

$$J_r = \frac{1}{\pi} \int_{R_{\text{mín}}}^{R} \left[2mE - \frac{p_\theta^2}{r^2} \right]^{\frac{1}{2}} dr. \tag{8.177}$$

Para hacer esta integral cambiamos de variable y usamos θ en vez de r (ver figura (8.48)). Entonces

$$r = \frac{R_{\text{mín}}}{\cos\theta}, \tag{8.178}$$

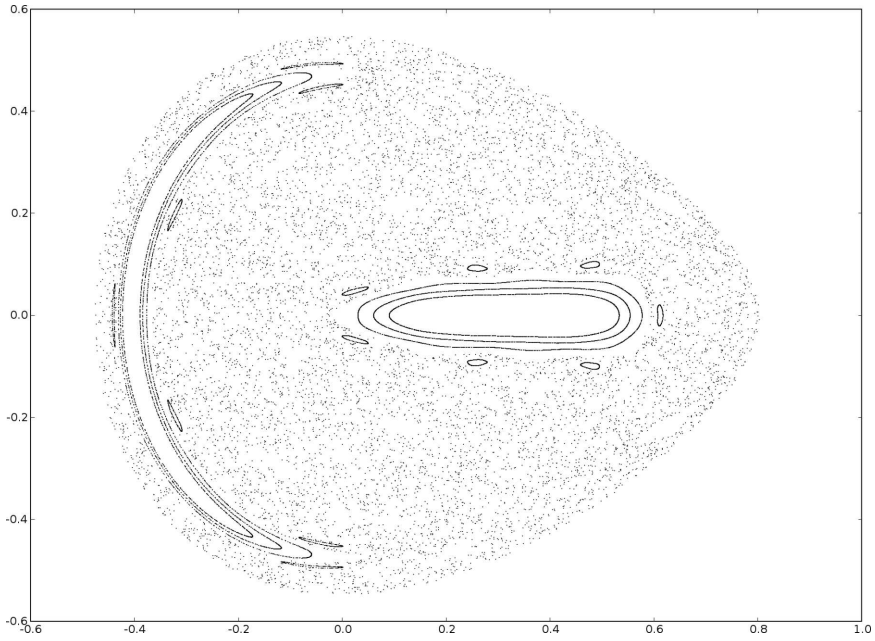


Figura 8.45: El sistema de Henon y Hailes, caso con energía total 0.15.

$$dr = R_{\text{mín}} \left(\frac{\sin\theta}{\cos^2\theta} \right) d\theta \tag{8.179}$$

y los valores de los extremos de integración serán $\theta = 0$ y $\theta = \phi$; entonces, tomando en consideración que

$$2mE = \frac{p_\theta^2}{R_{\text{mín}}^2}, \tag{8.180}$$

$$J_r = \frac{1}{\pi} R_{\text{mín}} \sqrt{2mE} \int_0^\phi \tan^2\theta d\theta = \frac{1}{\pi} R_{\text{mín}} \sqrt{2mE} \left[\tan\phi - \phi \right]. \tag{8.181}$$

Ahora como

$$\phi = \cos^{-1} \frac{p_\theta}{R\sqrt{2mE}}, \tag{8.182}$$

tendremos que, en términos de E y ϕ

$$J_\theta = R\sqrt{2mE}\cos\phi \tag{8.183}$$

y

$$J_r = \frac{R\sqrt{2mE}}{\pi} \left[\sin\phi - \phi\cos\phi \right]. \tag{8.184}$$

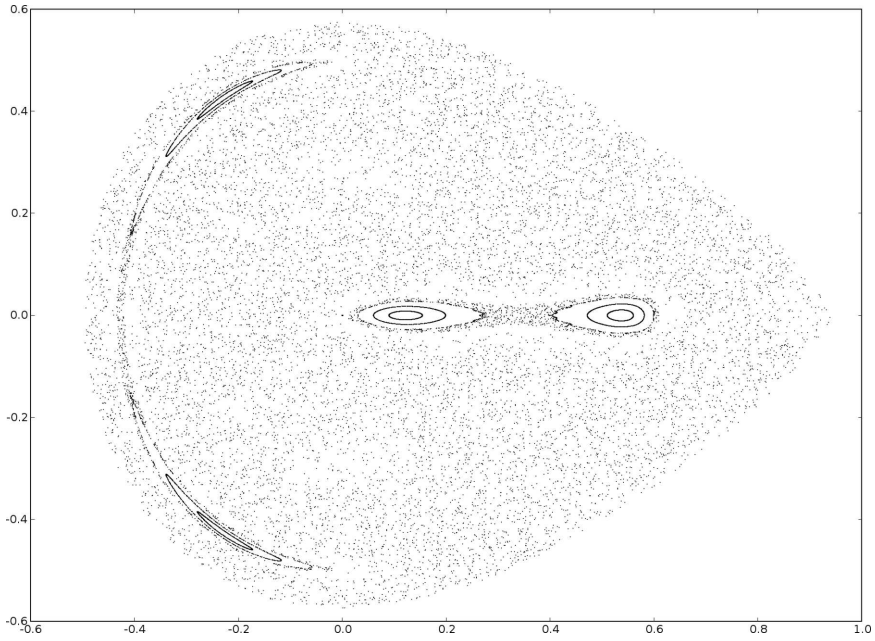


Figura 8.46: El sistema de Henon y Hailes, caso con energía total 0.166.

Esto mismo en términos de E y p_θ será

$$J_\theta = p_\theta \tag{8.185}$$

y

$$J_r = \frac{1}{\pi}\left[\left(2mER^2 - p_\theta^2\right)^{\frac{1}{2}} - p_\theta \cos^{-1}\frac{p_\theta}{R\sqrt{2mE}}\right]. \tag{8.186}$$

A partir de la última expresión podemos encontrar las frecuencias ω_r y ω_θ calculando las parciales de E con respecto a los impulsos correspondientes. La expresión ((8.186)) puede escribirse como

$$J_r = \frac{1}{\pi}\left[\left(2mER^2 - J_\theta^2\right)^{\frac{1}{2}} - J_\theta\phi\right], \tag{8.187}$$

calculamos las parciales con respecto a J_r y J_θ para obtener

$$1 = \frac{1}{\pi}\left[\frac{R^2 m}{\sqrt{2mER^2 - J_\theta^2}}\frac{\partial E}{\partial J_r} - J_\theta\frac{\partial \phi}{\partial J_r}\right] \tag{8.188}$$

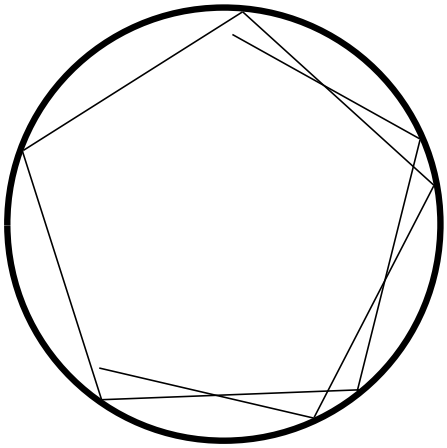


Figura 8.47: Una trayectoria típica en un billar circular. Podemos apreciar que estas trayectorias serán periódicas o cuasiperiódicas dependiendo de la conmensurabilidad del ángulo de incidencia con π.

y

$$0 = \frac{1}{\pi}\left[\frac{R^2 m\frac{\partial E}{\partial J_\theta} - J_\theta}{\sqrt{2mER^2 - J_\theta^2}} - \phi - J_\theta\frac{\partial \phi}{\partial J_\theta}\right]. \tag{8.189}$$

Las parciales de ϕ las podemos calcular de ((8.182))

$$\frac{\partial \phi}{\partial J_r} = \frac{J_\theta}{2E\sqrt{2mER^2 - J_\theta^2}}\frac{\partial E}{\partial J_r} \tag{8.190}$$

y

$$\frac{\partial \phi}{\partial J_\theta} = \frac{J_\theta}{2E\sqrt{2mER^2 - J_\theta^2}}\frac{\partial E}{\partial J_\theta} - \frac{1}{\sqrt{2mER^2 - J_\theta^2}}; \tag{8.191}$$

substituyendo encontramos que

$$\omega_r = \frac{\partial E}{\partial J_r} = \frac{\pi}{\sqrt{2mER^2 - J_\theta^2}} \tag{8.192}$$

y

$$\omega_\theta = \frac{\partial E}{\partial J_\theta} = \frac{\phi}{\sqrt{2mER^2 - J_\theta^2}}. \tag{8.193}$$

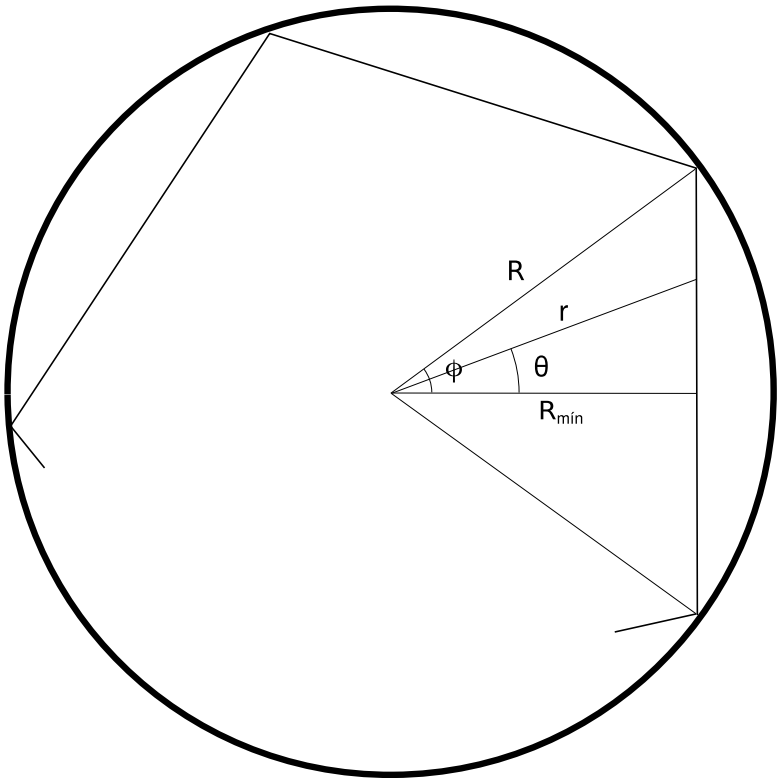


Figura 8.48:

De estas ecuaciones encontramos que

$$\phi = \frac{\omega_\theta}{\omega_r}\pi. \tag{8.194}$$

Esto nos dice que el movimiento será periódico o cuasiperiódico dependiendo de si ϕ es conmensurable o inconmensurable con π, pues esa es la conmensurabilidad entre ω_r y ω_θ.

Un sistema un poco más complejo consiste de un billar en el que una partícula esta constreñida a moverse en la región anular comprendida entre dos círculos concéntricos, uno de radio R y otro de radio ρ. Este billar, que se representa en la figura (8.49), tiene las mismas simetrías que el anterior, por lo que tiene las mismas integrales de movimiento y resulta integrable. Nuevamente tendremos trayectorias periódicas o cuasiperiódicas, pero en este caso la foliación por toros tiene una separatriz que la divide en dos partes correspondientes a dos tipos de movimiento distintos. Cuando las condiciones iniciales son tales que el radio mínimo

$$R_{\text{mín}} = \frac{p_\theta}{\sqrt{2mE}} \tag{8.195}$$

es mayor que ρ, las trayectorias son idénticas a las del billar anterior, mientras que en el caso contrario tendremos trayectorias que rebotan sucesivamente entre los dos círculos. Es importante hacer notar que la conservación del impulso angular hace que $\dot{\theta}$ sea siempre positiva o negativa, esto es, el sentido del giro de las trayectorias no cambia. La expresión de la conmensurabilidad de las frecuencias para el caso $\rho > R_{\text{mín}}$ son más complejas, pero es fácil encontrar geométricamente algunas trayectorias periódicas.

En la figura (8.50) se trazan las órbitas del mapeo del plano (r, p_r) con $\theta = 0$ en si mismo que se obtiene al comenzar una trayectoria en un punto de dicho plano, seguirla hasta que θ sea 2π, esto es hasta que la trayectoria vuelva a pasar por el plano y definir el mapeo por esa intersección. Es fácil ver que estas órbitas estarán sobre las curvas

$$p_r = \pm\sqrt{2mE - \frac{p_\theta^2}{r^2}} \tag{8.196}$$

cerradas por la recta vertical $r = R$ trazada entre la rama positiva y la negativa y por la recta vertical $r = \rho$ trazada entre estas dos ramas en el caso en que $R_{\text{mín}} > \rho$. Los toros en la variedad de energía constante se generan al dar estas curvas cerradas vueltas cuando la coordenada θ va de 0 a 2π. El paso de las trayectorias por las rectas verticales es instantáneo ya que corresponde a los choques con la frontera de la región.

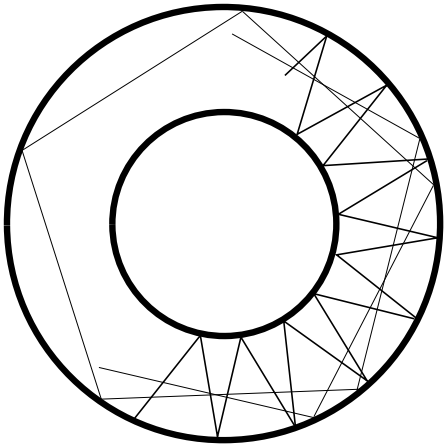


Figura 8.49: Billar compuesto por dos círculos concéntricos. Se ha trazado una órbita que no toca al círculo interno y otra que si lo hace.

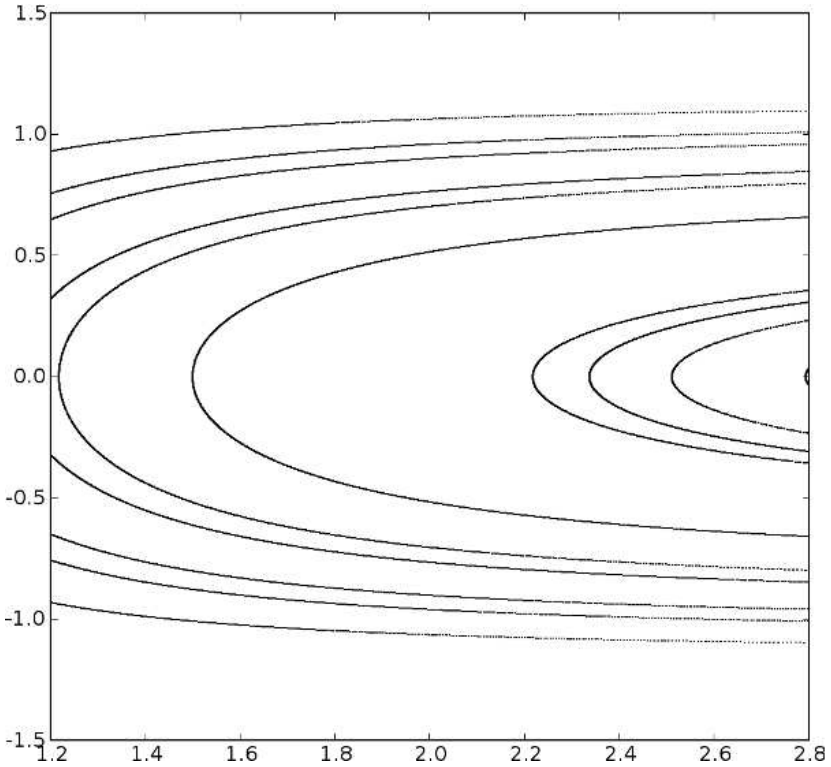


Figura 8.50: Los toros del billar concéntrico.

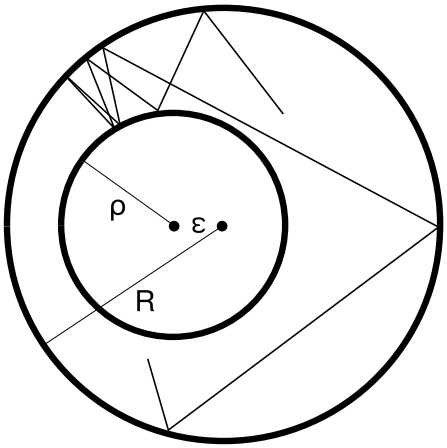


Figura 8.51: Un billar compuesto de dos círculos excéntricos. La distancia entre los centros es ϵ y se ha trazado una órbita.

Tomemos ahora este billar anular y desplacemos el centro del círculo interior una distancia ϵ, como se muestra en la figura (8.51).

Al hacer esto se pierde la simetría ante rotaciones y el impulso angular deja de ser una integral de movimiento; de hecho, como ocurre con la inmensa mayoría de los sistemas cercanos a uno integrable, el sistema se hace no integrable. La figura (8.51) muestra la existencia, para este sistema excéntrico, de trayectorias que invierten su sentido de rotación. Tendremos también trayectorias del tipo de las del primer billar cuando las condiciones iniciales sean tales que

$$R_{\text{mín}} = \frac{p_\theta}{\sqrt{2mE}} > \rho + \epsilon, \tag{8.197}$$

pues en ese caso el círculo interior queda dentro del círculo de radio $R_{\text{mín}}$, naturalmente esto será posible únicamente para sistemas en que $\rho + \epsilon < R$.

Para estudiar este sistema utilizaremos la iteración del mapeo del plano (r, p_r) en si mismo generado por las trayectorias cuando pasan por $\theta = 0$ con el signo de p_θ idéntico al de la p_θ inicial. En las figuras (8.52) a (8.57) se muestran varios casos correspondiente a un valor $R = 2.8$ del radio del círculo exterior y $\rho = 1.2$ para el radio del interior, podemos ver claramente la persistencia de algunos de los toros a medida que el sistema se aleja del integrable según lo establecido por el teorema de KAM y podemos ver también la presencia de órbitas caóticas. La variable ϵ se ha tomado dentro de los valores cero, correspondiente al caso integrable, y 1.6, caso en que los dos círculos se tocan y ya no existen órbitas que no chocan con el círculo

interior.

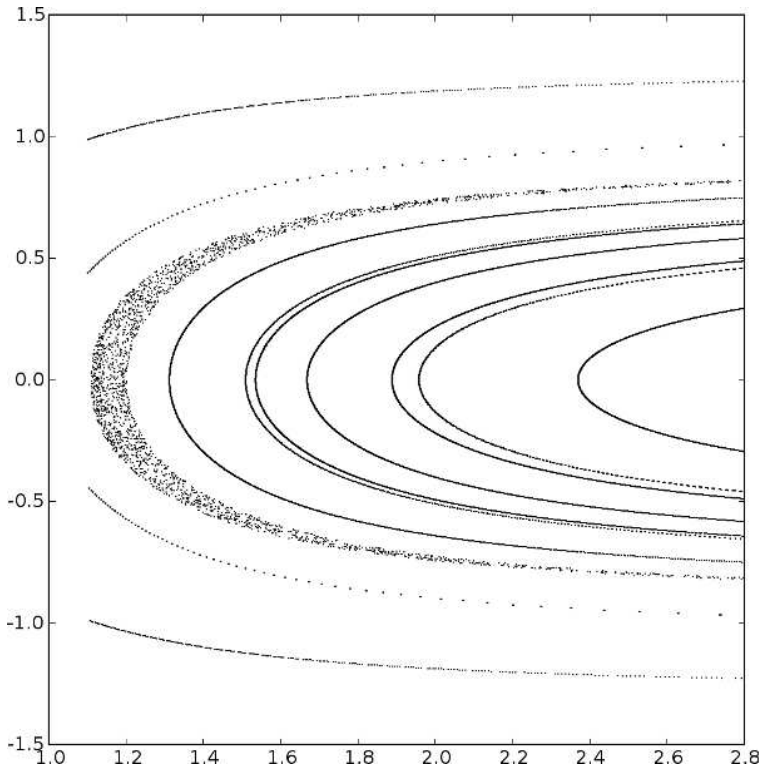


Figura 8.52: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 0.1$.

8.5.4. El péndulo doble.

El péndulo doble nos sirve como un último ejemplo. En el sistema de coordenadas mostrado en la figura (8.58) y para el caso en que las dos masas y las dos longitudes son iguales y unitarias la hamiltoniana es

$$H = \frac{1}{2}\left[\frac{p_1^2 + 2p_2^2 - 2p_1 p_2 \cos(\theta_1 - \theta_2)}{1 + \text{sen}^2(\theta_1 - \theta_2)}\right] - g\left(2\cos\theta_1 + \cos\theta_2\right). \qquad (8.198)$$

Esta forma de la hamiltoniana nos permite ilustrar los fenómenos que ocurren cerca de los sistemas integrables pues para el valor cero de g el sistema es integrable y podemos utilizar a g como un parámetro de perturbación, naturalmente g tiene un valor físico, pero como ilustración el paso del sistema integrable con $g = 0$ al sistema físico es interesante.

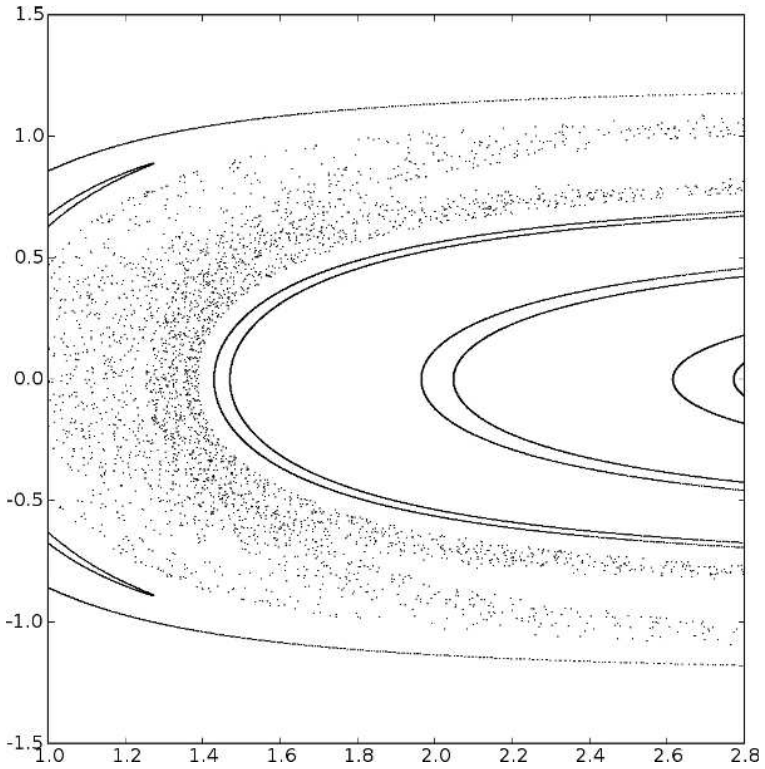


Figura 8.53: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 0.2$.

Las ecuaciones de movimiento son

$$\dot{\theta}_1 = \frac{p_1 - p_2 \cos(\theta_1 - \theta_2)}{1 + \sin^2(\theta_1 - \theta_2)}, \tag{8.199}$$

$$\dot{\theta}_2 = \frac{2p_2 - p_1 \cos(\theta_1 - \theta_2)}{1 + \mathrm{sen}^2(\theta_1 - \theta_2)}, \tag{8.200}$$

$$\dot{p}_1 = \frac{\mathrm{sen}(\theta_1 - \theta_2)\cos(\theta_1 - \theta_2)(p_1^2 + 2p_2^2 - 2p_1 p_2 \cos(\theta_1 - \theta_2)}{1 + \mathrm{sen}^2(\theta_1 - \theta_2)} - \tag{8.201}$$

$$- \frac{p_1 p_2 \mathrm{sen}(\theta_1 - \theta_2)}{1 + \mathrm{sen}^2(\theta_1 - \theta_2)} - 2g\mathrm{sen}\theta_1, \tag{8.202}$$

$$\dot{p}_2 = -\frac{\mathrm{sen}(\theta_1 - \theta_2)\cos(\theta_1 - \theta_2)(p_1^2 + 2p_2^2 - 2p_1 p_2 \cos(\theta_1 - \theta_2)}{1 + \mathrm{sen}^2(\theta_1 - \theta_2)} + \tag{8.203}$$

$$+ \frac{p_1 p_2 \mathrm{sen}(\theta_1 - \theta_2)}{1 + \mathrm{sen}^2(\theta_1 - \theta_2)} - g\mathrm{sen}\theta_2. \tag{8.204}$$

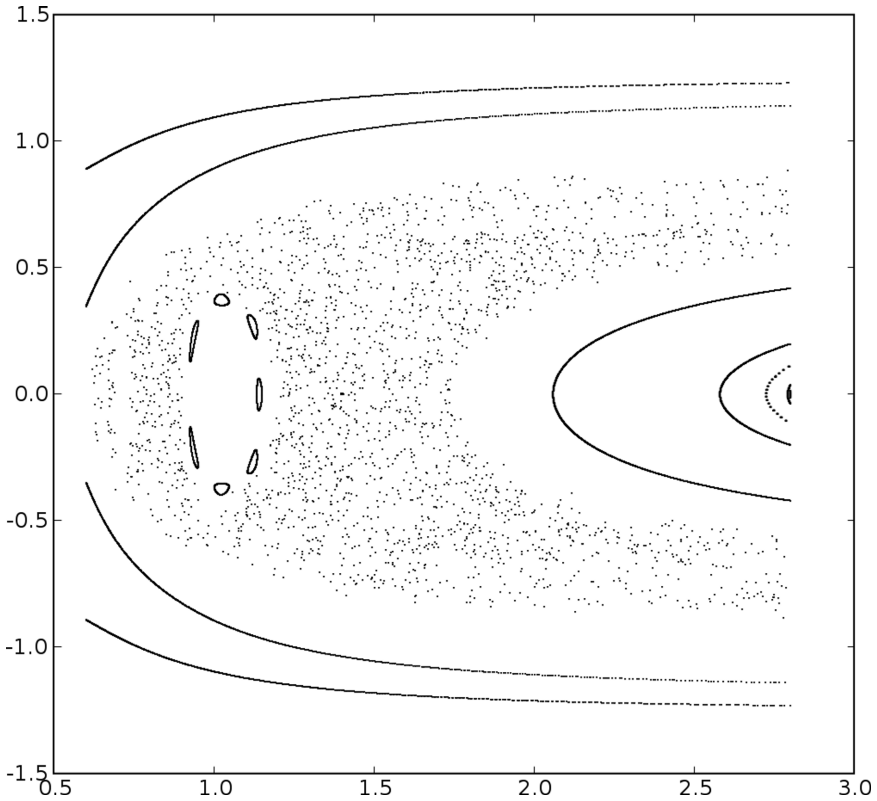


Figura 8.54: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 0.6$.

Por simple inspección de estas ecuaciones encontramos que en el caso integrable ($g = 0$) la cantidad $(p_1 + p_2)$ es una integral de movimiento que con la hamiltoniana, que en este caso es la energía total, conforman un sistema independiente de integrales de movimiento en involución. De hecho podemos pasar a un nuevo sistema canónico de coordenadas dado por la transformación canónica

$$\Theta_1 = \theta_1 - \theta_2, \tag{8.205}$$

$$\Theta_2 = \theta_1 + \theta_2, \tag{8.206}$$

$$P_1 = \frac{1}{2}(p_1 - p_2), \tag{8.207}$$

$$P_2 = \frac{1}{2}(p_1 + p_2), \tag{8.208}$$

en el que, para $g = 0$, Θ_2 es ignorable y, por tanto, P_2 una integral de movimiento.

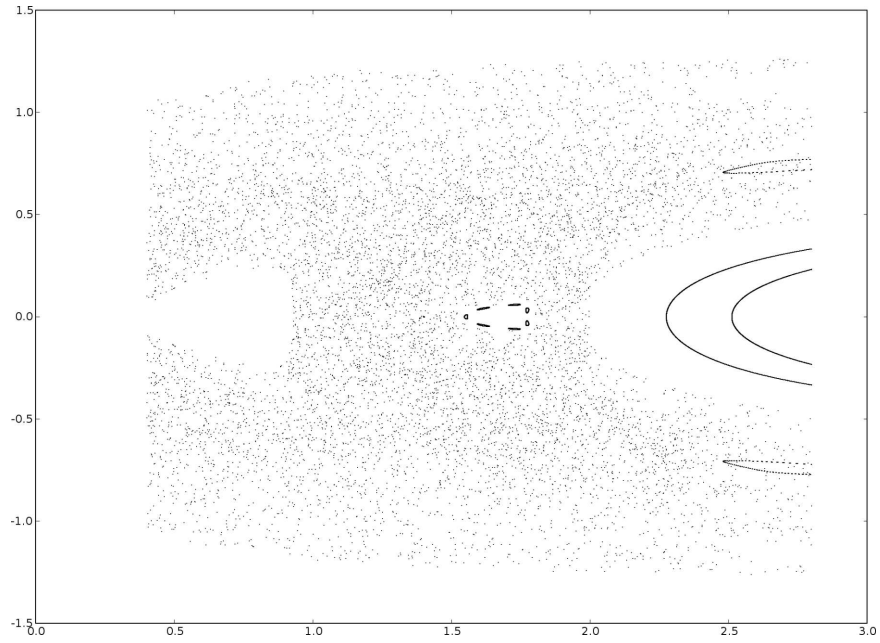


Figura 8.55: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 0.8$.

En las figuras se presentan soluciones numéricas para distintos valores de la energía total y de g, según se indica en cada una de ellas. En la gráfica superior tenemos la proyección de la trayectoria en el espacio fase sobre el 'plano coordenado' (Θ_1, P_1) trazada por dos mil puntos obtenidos cada intervalo unitario del tiempo; se superponen en la misma gráfica cinco trayectorias con la misma energía pero con condiciones iniciales establecidas aleatoriamente. En la gráfica inferior se presenta de manera continua la curva trazada por la segunda masa sobre el plano del espacio físico, en este caso para no complicar la figura se presenta sólo una solución desarrollada por un tiempo de veinte unidades.

En los casos en que $g = 0$ las trayectorias están sobre toros de dos dimensiones, lo que se refleja en las gráficas superiores. Debemos desde luego tomar en consideración que el espacio fase en este caso no es R^4, es el producto cartesiano del toro de configuración con R^2, razón por la cual en la proyección (Θ_1, P_1), es difícil evitar cruzamientos; desde luego vemos directamente curvas cerradas y curvas que se cierran cuando tomamos en consideración que (Θ_1, P_1) no es un plano, por lo que hay que identificar el borde derecho de la gráfica con el izquierdo.

Las figuras muestran claramente como se pasa del caso integrable al caó-

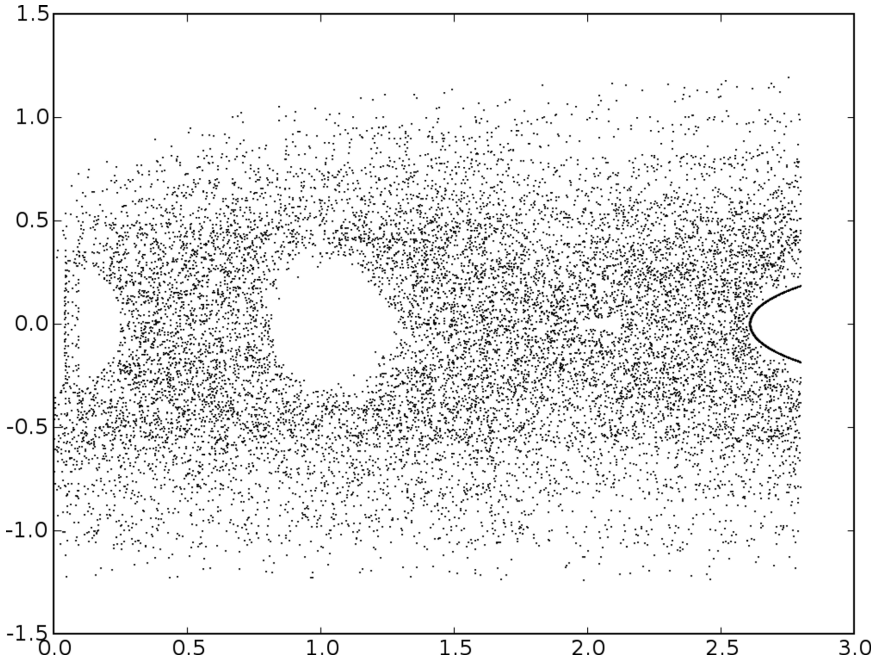


Figura 8.56: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 1.4$.

tico; para el caso realista en que $g = 9.8$[14] vemos la enorme complejidad de las trayectorias y del movimiento.

[14] Este caso representa un péndulo doble de longitudes iguales a un metro y masas iguales a un kilogramo.

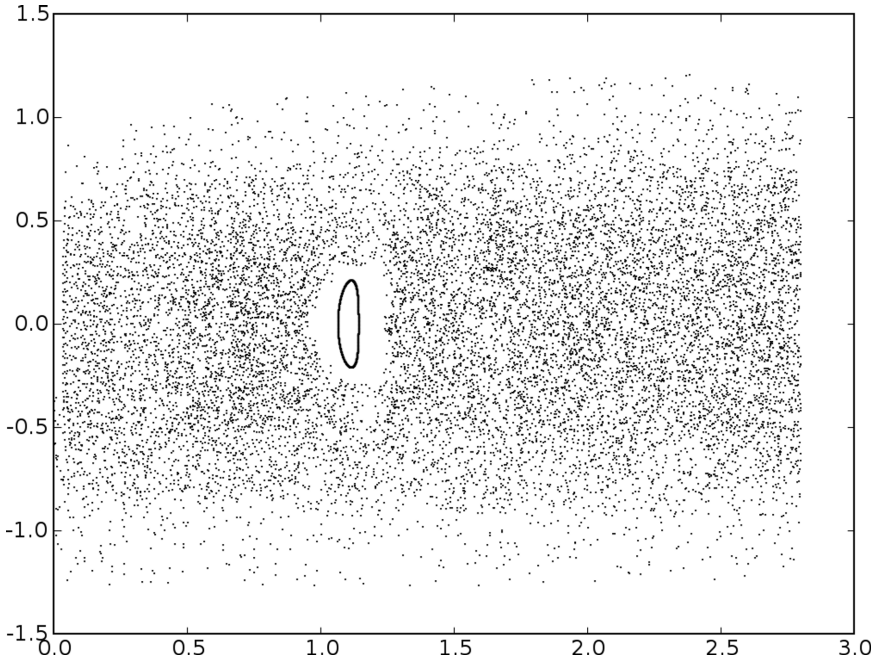
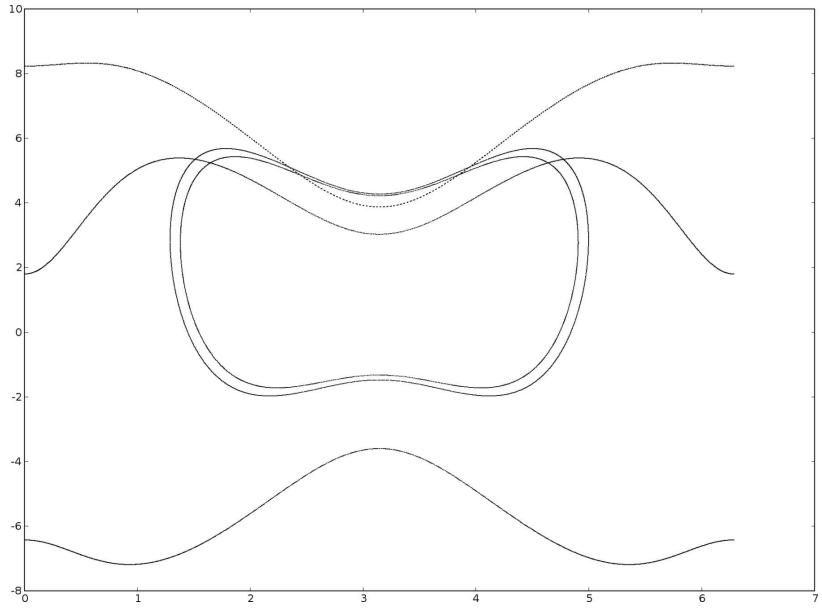
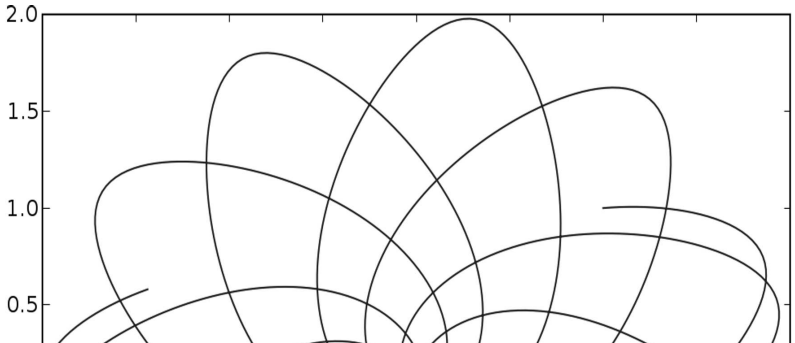


Figura 8.57: Los toros del billar excéntrico. $R = 2.8$, $\rho = 1.2$ y $\epsilon = 1.6$.





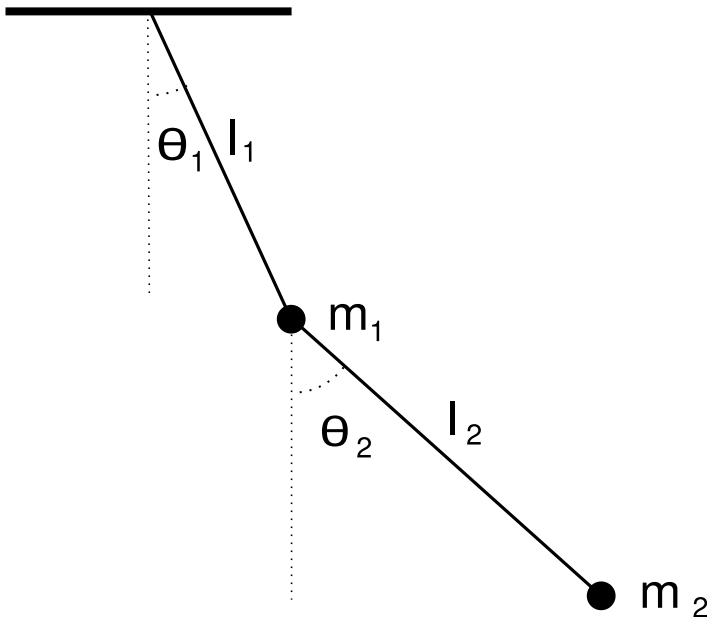
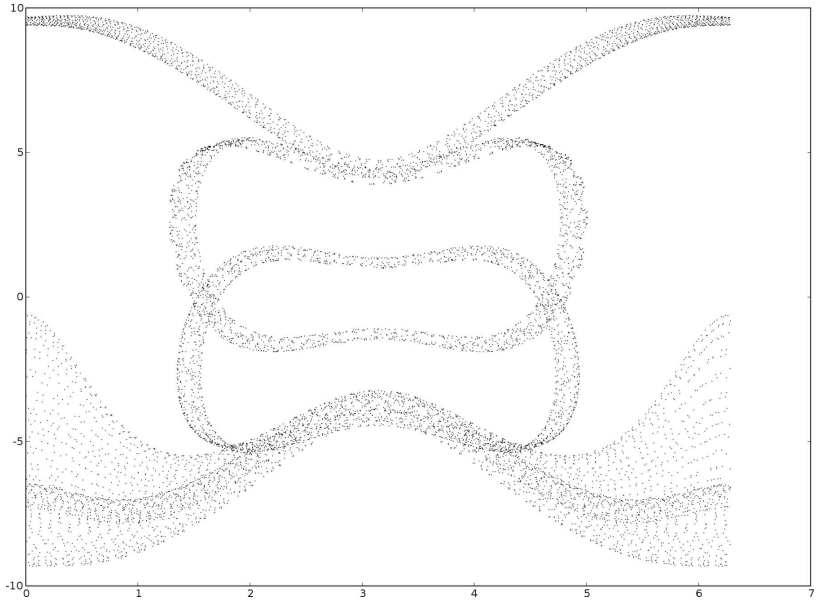
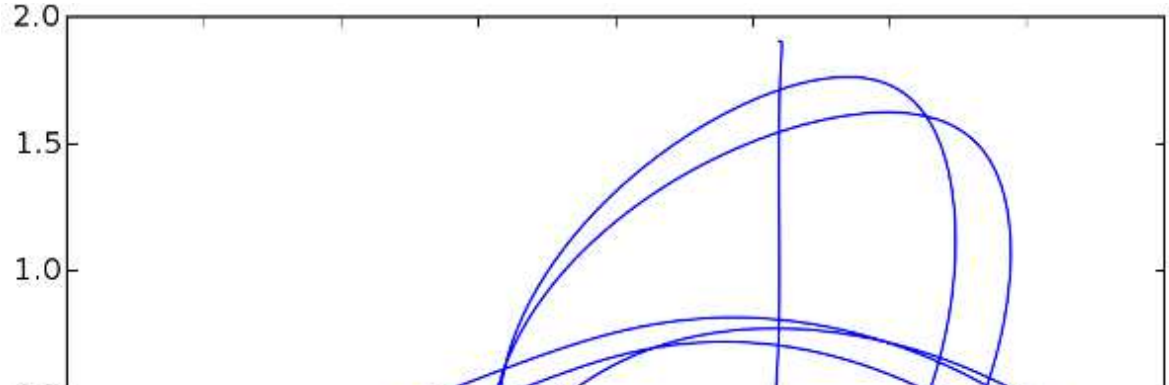


Figura 8.58: El péndulo doble.





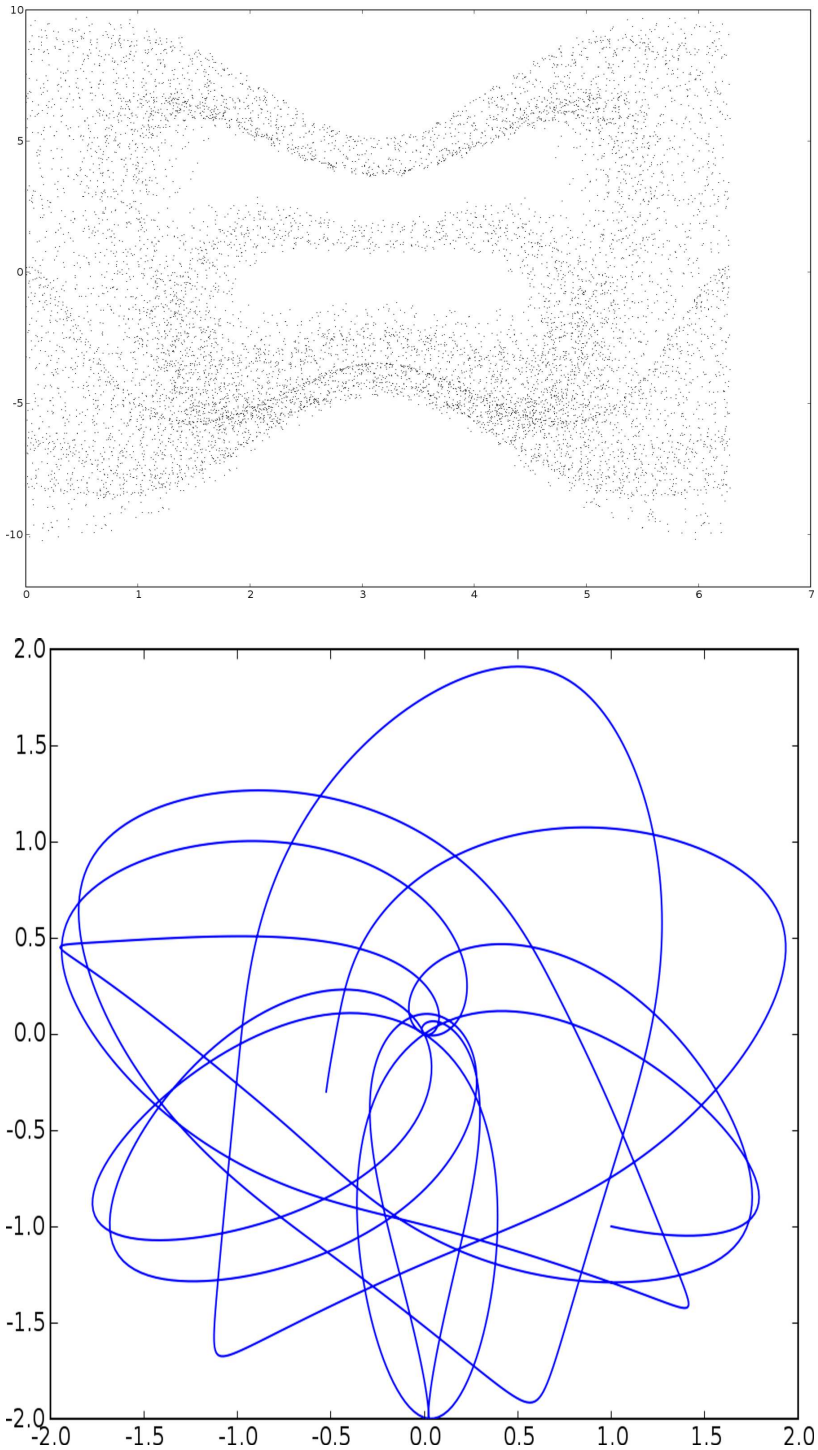


Figura 8.61: El péndulo doble para una energía total de 10.1 unidades y un valor $g = 1.0$.

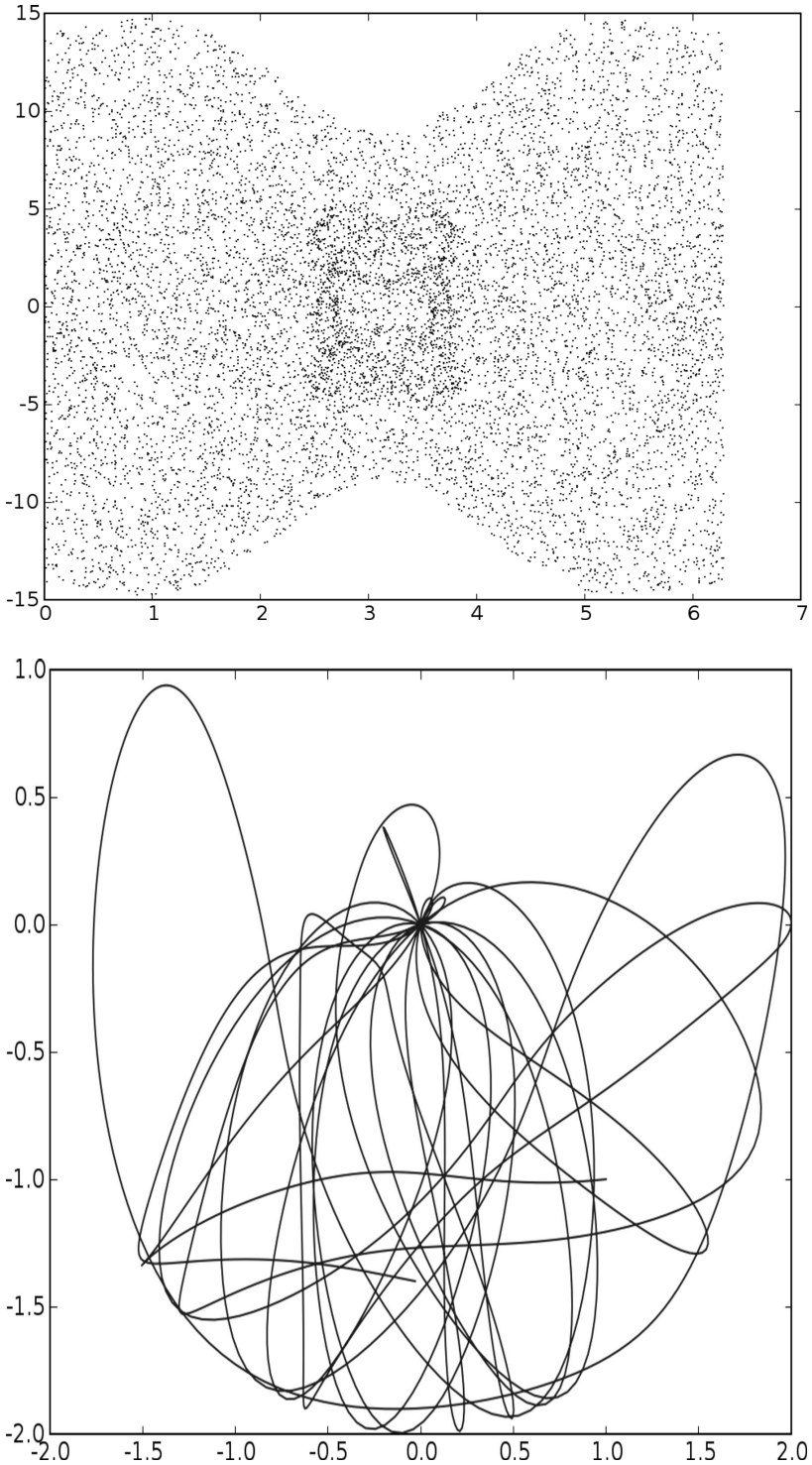


Figura 8.62: El péndulo doble para una energía total de 10.1 unidades y un valor $g = 9.8$.

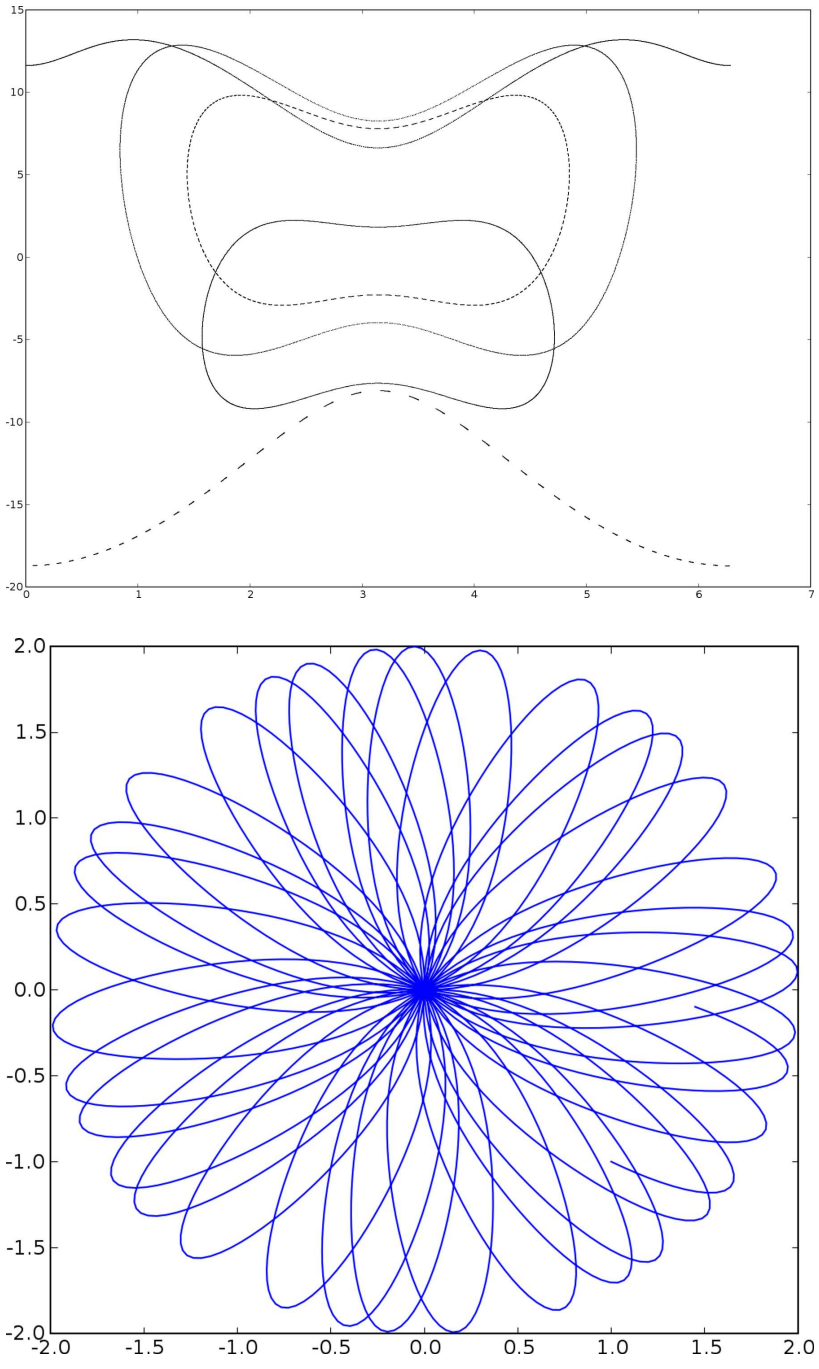


Figura 8.63: El péndulo doble para una energía total de 35.0 unidades y un valor $g = 0.0$.

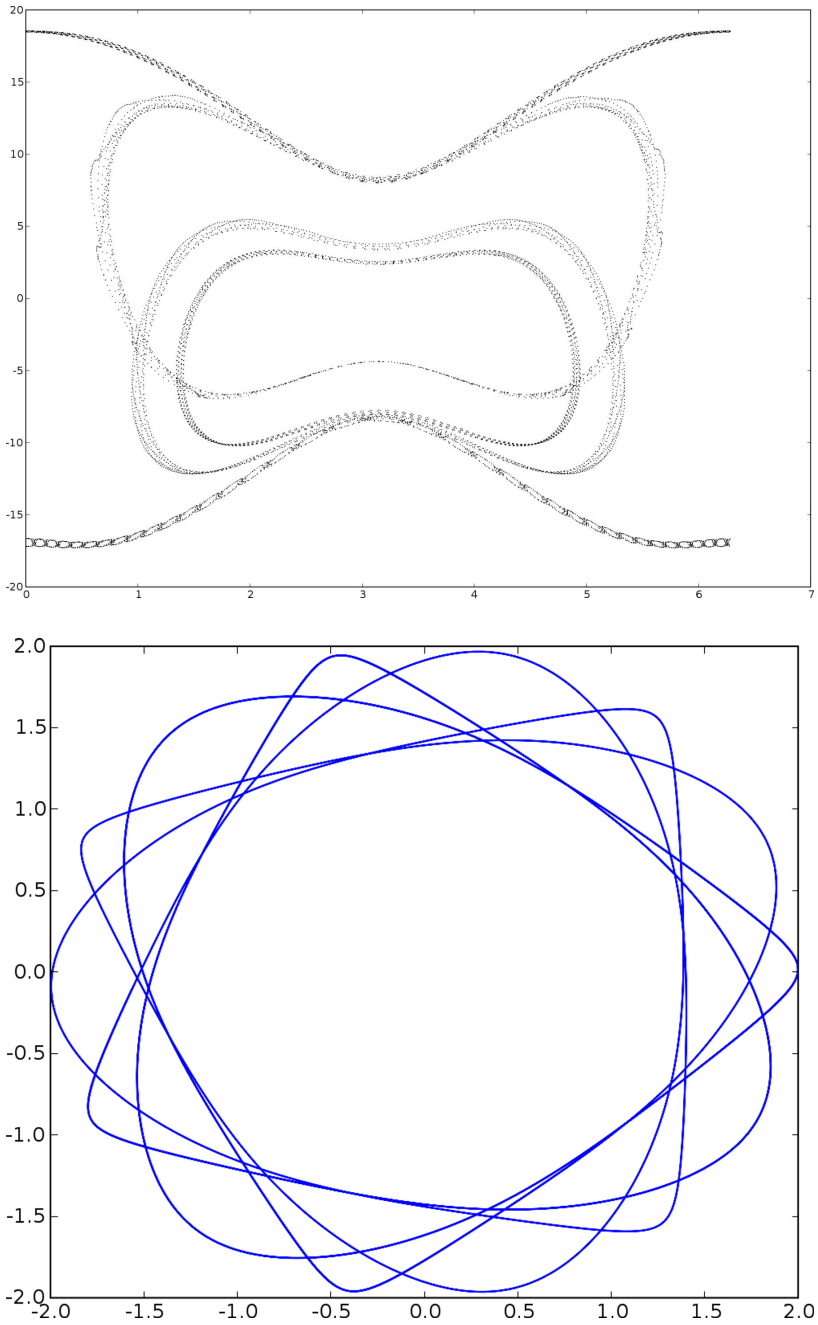


Figura 8.64: El péndulo doble para una energía total de 35.0 unidades y un valor $g = 0.5$.

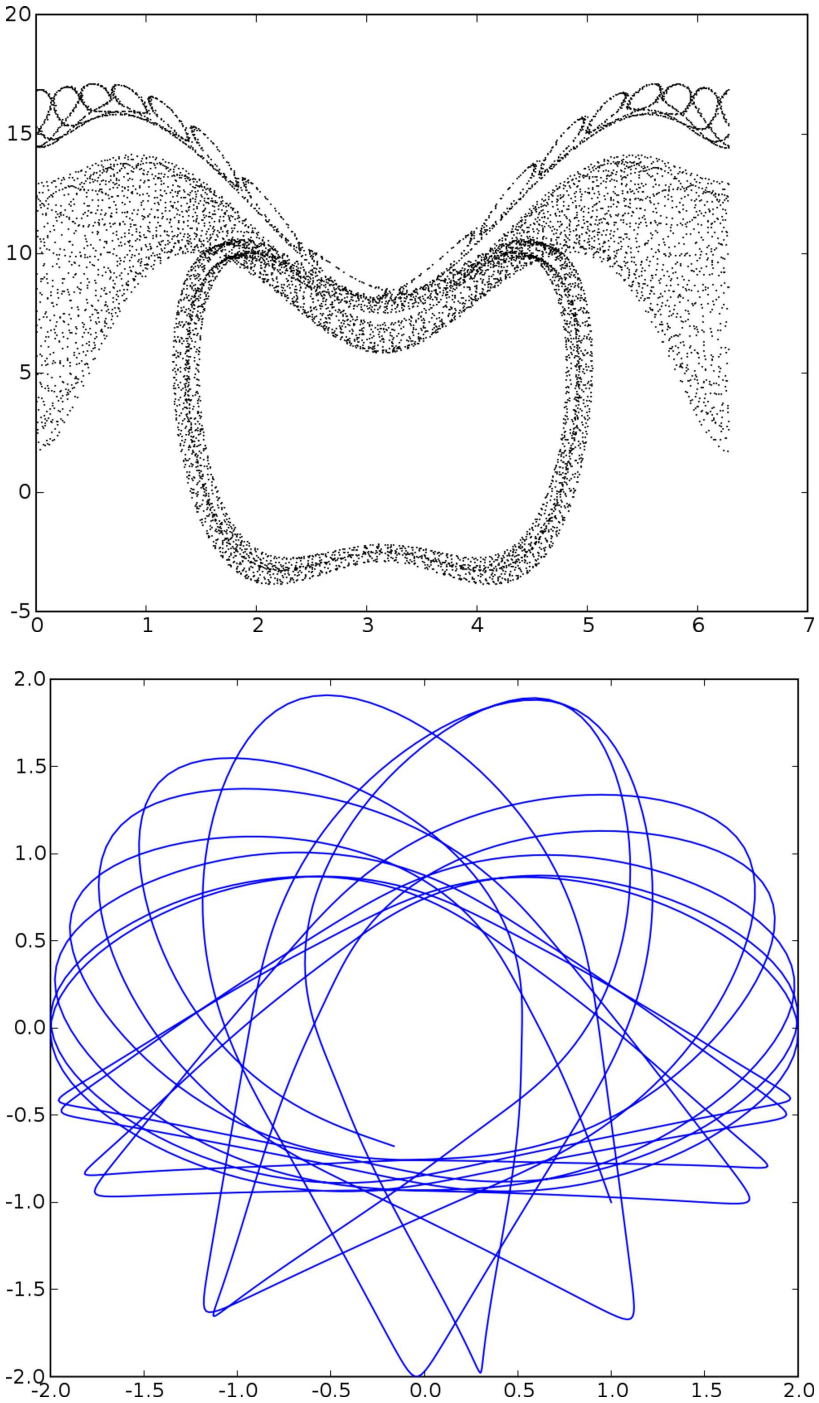


Figura 8.65: El péndulo doble para una energía total de 35.0 unidades y un valor $g = 1.0$.

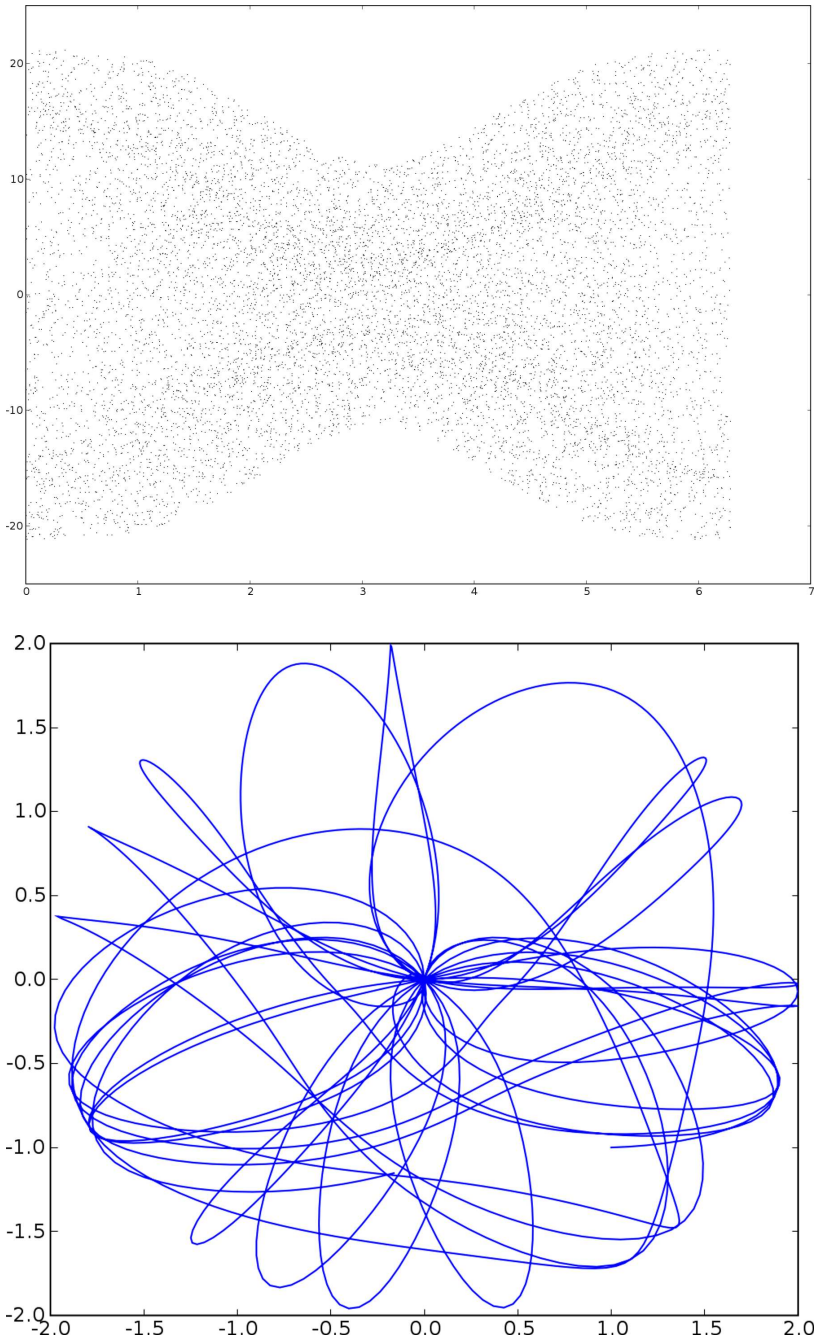


Figura 8.66: El péndulo doble para una energía total de 35.0 unidades y un valor $g = 9.8$.

8.6. Caracterización del caos.

En los ejemplos anteriores hemos utilizado la apariencia de las figuras para detectar el comportamiento periódico, cuasiperiódico o caótico de diversos sistemas, esto, si bien es muy eficaz, no es riguroso, no permite hacer estudios del caos o su caracterización y es difícilmente aplicable a casos de más de dos grados de libertad. Cuando se trata de sistemas de pocos grados de libertad, como los que nos han servido de ejemplo, la distinción es clara, pero cuando tenemos más grados de libertad resulta muy difícil distinguir entre las trayectorias cuasiperiódicas y las caóticas por medio de la apariencia de la figura de alguna proyección de la trayectoria o por la gráfica temporal de alguna coordenada; incluso se llegó a pensar que los fenómenos físicos en los que se presenta un comportamiento irregular, como la turbulencia, se trataban de fenómenos cuasiperiódicos en los que participaban muchas frecuencias, lo que en un caso mecánico correspondería a muchos grados de libertad. En realidad cuando tenemos una secuencia de datos experimentales o provenientes de soluciones numéricas y se tienen muchos grados de libertad es sumamente difícil distinguir si se trata de un fenómeno caótico o a uno cuasiperiódico, incluso las técnicas que hoy se tienen no proporcionan, en muchos casos, una total certidumbre; únicamente cuando tenemos información adicional, por ejemplo conocemos el número de grados de libertad, dichas técnicas se hacen plenamente útiles. Pasemos pues a estudiar algunos de los métodos más utilizados en el estudio de las trayectorias caóticas.

El problema de distinguir entre trayectorias caótica, periódicas o cuasiperiódicas es análogo a los problemas que se enfrentan en el análisis de señales, por lo que algunas de las técnicas de ese campo pueden adaptarse al estudio de trayectorias de sistemas dinámicos o de datos experimentales. La más básica y utilizada de las técnicas en el caso de las señales es el análisis de Fourier. Una señal, digamos la que captamos con un receptor de radio, es una función del tiempo $s(t)$, que puede ser la intensidad o la amplitud. Esta función puede ser periódica, cuasiperiódica o no tener características de periodicidad, esto es, ser aperiódica o incluso totalmente aleatoria, pero en todo caso siempre la podemos analizar en términos de la intensidad, amplitud, energía o potencia que presenta en función de la frecuencia de las componentes armónicas en que se puede descomponer. Esto es, en vez de estudiar $s(t)$, estudiamos su transformada de Fourier $S(f)$ o mejor aun el llamado espectro de potencias dado por

$$P(f) = |S(f)|^2 , \tag{8.209}$$

que tiene la ventaja de ser siempre real y positivo. Si se trata de una señal

periódica y armónica, $P(f)$ será una delta en la frecuencia correspondiente, si fuese periódica, pero no harmónica, $P(f)$ presentará picos, de diversas alturas, en la frecuencia correspondiente y sus armónicos. Si se trata de una señal cuasiperiódica, $P(t)$ presentará picos en cada una de las frecuencias básicas, en sus armónicos y en las combinaciones por enteros de esas frecuencias, esto es si tenemos dos frecuencias, f_1 y f_2, tendremos picos en esas frecuencias y sus combinaciones de la forma $mf_1 + nf_2$ donde m y n son números enteros, la altura o valor de esos picos, en general, decae exponencialmente a medida que m y n crecen, por lo que sólo suelen observarse los correspondientes a valores pequeños de m y n. Si la señal es aperiódica $P(f)$ será continua o lo que es lo mismo presentará picos en todos los valores de f, naturalmente si la señal se aproxima a tener algunas periodicidades, $P(f)$ presentará algunos máximos pronunciados dentro de ese continuo, pero si no presenta ese tipo de regularidades se tendrá una variación sin grandes máximos. En general no se cuenta con la señal $s(t)$ para t de menos a más infinito, se tiene únicamente para un cierto intervalo finito de valores de t, esto obliga a calcular la transformada de Fourier únicamente para ese intervalo, o mejor dicho, para una función de t periódica cuyo período sea ese intervalo temporal. Esto introduce una frecuencia artificial que debe siempre tomarse en cuenta.

Esta técnica es plenamente aplicable para el caso de las soluciones de las ecuaciones diferenciables y nos permite explorar si presentan un comportamiento periódico, cuasiperiódico o aperiódico, al que en este campo denominamos caótico. Podemos aplicar la técnica a cualquier variable de interés que dependa del tiempo vía la trayectoria de la ecuación diferencial, en muchos casos se utiliza alguna de las coordenadas en el espacio fase o una combinación apropiada de ellas. Podemos también aplicarla a datos experimentales para orientar el estudio de algún sistema de interés.

En adición a la limitante de no contar con un desarrollo temporal infinito, en la mayor parte de las ocasiones, no contamos con la continuidad temporal. En el caso de las trayectorias de una ecuación diferencial tenemos soluciones numéricas, las que siempre consisten de un conjunto de valores de las coordenadas o de una variable de interés para una secuencia discreta y finita de valores del tiempo. En el caso experimental ocurre lo mismo, en general se tiene los datos tomados a valores discretos del tiempo y rara vez podemos actuar sobre la continuidad temporal, salvo aquellos casos en que el sistema experimental se acopla analógicamente con la electrónica que genera la transformada de Fourier. La mayor parte de las veces contamos únicamente con secuencias discretas y finitas de valores para tiempos conformados, a partir de un cierto tiempo inicial, por un incremento, generalmente constante, que

genera una secuencia temporal

$$t_0, t_1 = t_0 + \Delta t, t_2 = t_0 + 2\Delta t, \ldots, t_{n-1} = t_0 + (n-1)\Delta t. \tag{8.210}$$

Ante esta situación debemos utilizar la llamada transformada discreta de Fourier.

Dados los valores x_j de cierta variable $x(t)$ para los tiempos $t_j = t_0 + j\Delta t$, se define la transformada discreta de Fourier por

$$X_k = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} x_j e^{-i\frac{2\pi jk}{n}} \qquad k = 0, \ldots n-1. \tag{8.211}$$

Esta transformada discreta tiene una inversa dada por

$$x_j = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} X_k e^{i\frac{2\pi jk}{n}} \qquad j = 0, \ldots n-1. \tag{8.212}$$

Naturalmente cuando todas las x_j son todas reales se presenta la simetría

$$X_k = X^*_{n-k}, \tag{8.213}$$

donde el asterisco denota conjugación compleja.

Al igual que en el caso continuo la transformada discreta nos permite detectar propiedades de periodicidad de la secuencia x_j, claro está, cuando esta secuencia proviene de una variable continua, tendremos únicamente una aproximación al espectro que proporcionaría la transformada integral; en particular destacamos el hecho de no poderse detectar frecuencias mayores a $1/\Delta t$, ya que al discretizar el tiempo se pierde la información correspondiente. El valor de X_k se relaciona con la presencia de una frecuencia $1/k\Delta t$ y como tenemos únicamete una secuencia finita se ha introducido un período $n\Delta t$ o frecuencia $1/n\Delta t$ que es artificial, pues supone que la secuencia es tal que, con l un entero cualquiera, $x_{j+ln} = x_j$, lo que implica que $X_{k+ln} = X_k$.

Los valores de la transformada X_k son números complejos que en el caso de una señal representan una amplitud compleja, debido a esto se utiliza el cuadrado de su módulo en la busqueda de periodicidades. Esto ofrece la ventaja de trabajar con números reales positivos que se corresponden con la intensidad de la componente de la frecuencia correspondiente, razón por la que a la sucesión $|X_k|^2$ se le llama el espectro de potencias de la señal, término que se conserva para otros casos como el que nos interesa aquí. Es necesario hacer ver que el espectro de potencias, debido a la simetría ((8.213)), contiene únicamente la mitad de la información que contiene la

señal, se ha perdido la correspondiente a las fases de cada componente; esto, para el caso de un análisis de periodicidades, no es muy importante.

Cuando estudiamos una señal o variable de un sistema continuo utilizando una muestra discreta encontramos una dificultad adicional. Puesto que las frecuencias del sistema continuo no son de la forma $1/k\Delta t$ con k un entero, únicas contenidas en el análisis discreto, en el espectro de potencias tendremos que, la presencia de una frecuencia en el sistema continuo, genera un valor grande del espectro en la frecuencia discreta más cercana a ella y valores más pequeños en sus vecinas; esto se aplica también a las combinaciones enteras de frecuencias que se generan en el caso cuasiperiódico. Así el espectro de potencias para los casos periódicos o cuasiperiódicos no tiene valores distintos de cero en algunas de las frecuencias discretas y cero en las demás, presenta máximos en ciertas frecuencias (las más cercanas a las reales, sus armónicos o sus combinaciones enteras) y un fondo de valores pequeños en las demás. Afortunadamente podemos controlar esto haciendo Δt lo más pequeño posible, de esta forma la aproximación entera a la frecuencia real será mejor y la separación entre los máximos del espectro y el fondo más nítida, esto permite en muchos casos encontrar las periodicidades presentes en la señal; sin embargo para el estudio de posibles sistemas con características aperiódicas o caóticas, tan importante para nosotros, esto puede generar confusión, en especial en el caso experimental en el que se hace difícil distinguir entre un fenómeno determinista caótico, uno cuasiperiódico con muchas frecuencias o grados de libertad o uno en el que la intromisión inevitable del ruido se confunde con el caos determinista.

En las figuras (8.67), (8.68) y (8.69) se utiliza al péndulo doble para ilustrar el uso del espectro de potencias para caracterizar el movimiento. En la gráfica de superior se muestra la evolución temporal de la posición de la segunda masa en el plano físico de un péndulo doble y en la inferior el espectro de potencias de una secuencia compuesta de 2000 valores de la coordenada p_2 tomados cada 0.1 unidades de tiempo. Esto se hace para el caso integrable ($g = 0$), para el caso en $g = 0.2$ y para el caso físico ($g = 9.8$).

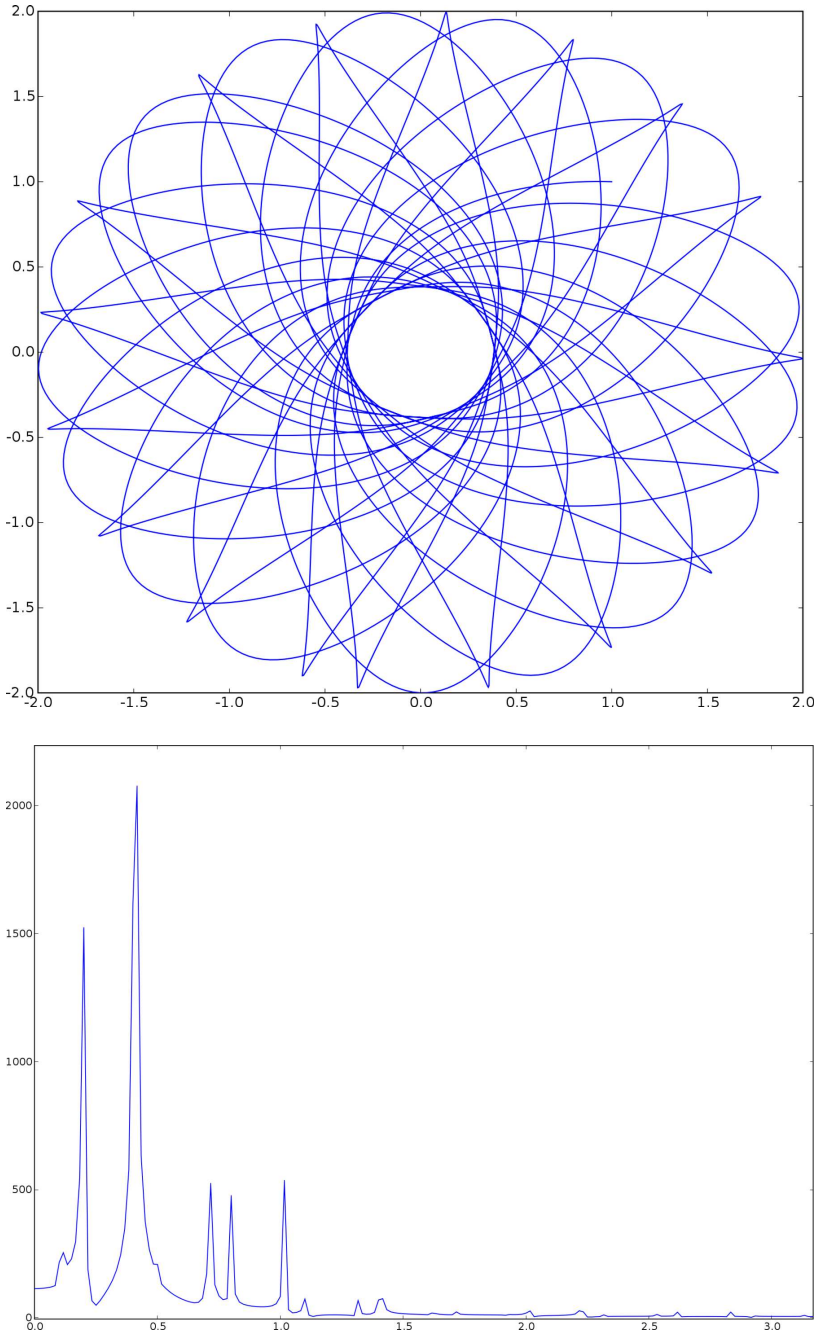


Figura 8.67: El espectro de Fourier para el péndulo doble con energía total de 10.0 unidades en el caso integrable, $g = 0$.

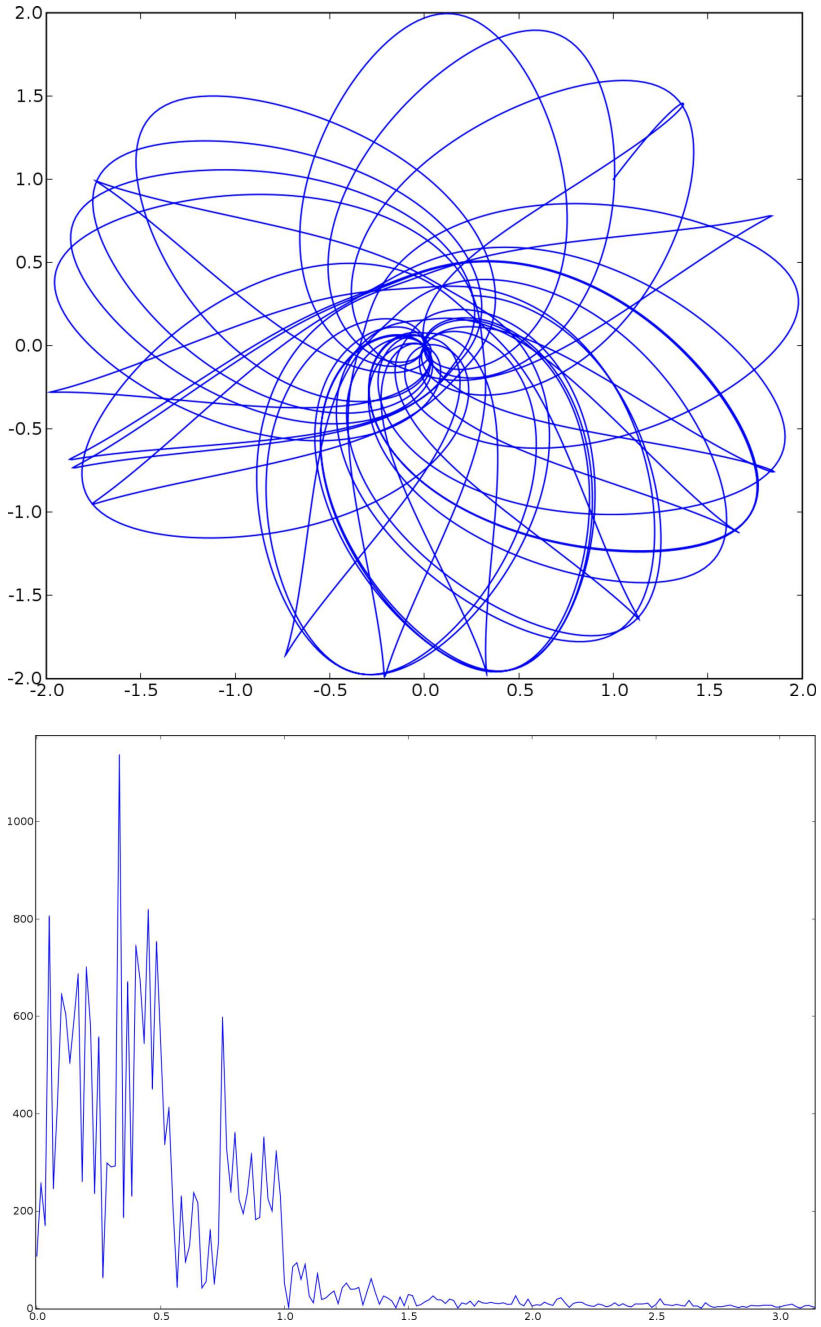


Figura 8.68: El espectro de potencias para el péndulo doble con energía total de 10.0 unidades cuando $g = 0.2$.

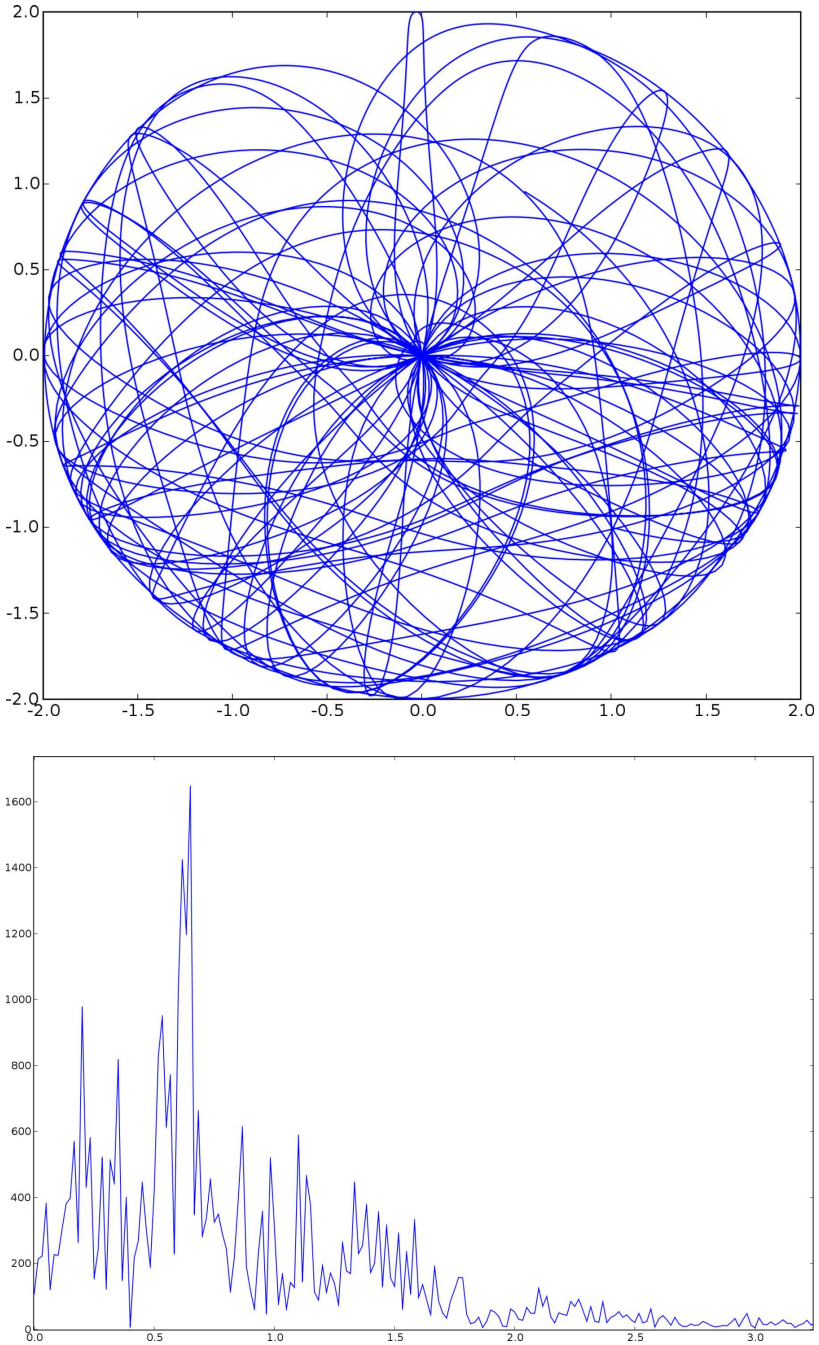


Figura 8.69: El espectro de potencias para una trayectoria del péndulo doble con energía total de 10.0 unidades cuando $g = 9.8$.

Podemos observar de las figuras como es posible detectar el caos a partir del espectro de potencias. El caso cuasiperiódico, ((8.67)), en principio, únicamente debe presentar valores distintos de cero en las frecuencias resultantes de combinaciones con coeficientes enteros de las dos frecuencias básicas del movimiento y su altura debe decaer exponencialmente con la magnitud de dichos coeficientes; las aproximaciones numéricas hacen que las lineas se ensanchen y el espectro puede llegar a confundirse con uno continuo parecido al correspondiente a un sistema caótico. En los casos en que conocemos el número de grados de libertad y estos son pocos, esta confusión es fácilmente eliminada estudiando las relaciones entre los picos más prominentes, en nuestro ejemplo esto se ve fácilmente.

Otra fuente de confusión cuando desconocemos el origen de una secuencia de datos experimentales son las dificultades para distinguir entre el espectro de un sistema determinista caótico y el producido por la presencia de elementos genuinamente aleatorios que presentan un espectro continuo. En el caso de los sistemas mecánicos esto ocurre pocas veces ya que en general conocemos la procedencia de los datos y la presencia de la energía o la hamiltoniana como integrales de movimiento o de la conservación del volumen en el espacio fase nos pueden orientar para distinguir entre sistemas mecánicos caóticos y datos provenientes de sistemas aleatorios.

En otra dirección tenemos que una de las características fundamentales del caos es la alta sensibilidad de las trayectorias al cambio en las condiciones iniciales, en general decimos que hay alta sensibilidad si trayectorias cercanas divergen exponencialmente al transcurrir el tiempo. Esto nos lleva a técnicas para medir esta divergencia y a criterios derivados de estas técnicas para caracterizar el caos.

Una medida de la sensibilidad a condiciones iniciales la encontramos en los llamados coeficientes de Liapunov. Tomemos un sitema de n dimensiones, no necesariamente mecánico,

$$\dot{x} = f(x) \tag{8.214}$$

y una trayectoria $x_0(t)$ correspondiente a unas condiciones iniciales $x_0 = x_0(0)$. Una trayectoria cercana será $x(t)$ con condiciones iniciales $x(0) = x_0 + \delta(0)$, donde δ está compuesta de n cantidades $\delta_i(0)$. Las soluciones presentarán la relación

$$x(t) = x_0(t) + \delta(t), \tag{8.215}$$

y, a primer orden, tendremos

$$\frac{d}{dt}\delta_i = \sum_j \left(\frac{\partial f_i}{\partial x_j}\bigg|_{x_0(t)} \right) \delta_j. \tag{8.216}$$

Definiendo la norma $s(t)$ por medio de

$$s(t) = \sqrt{\sum_i \delta_i^2(t)}, \tag{8.217}$$

podemos cuantificar la separación entre las dos soluciones usando la cantidad

$$\sigma = \lim_{t\to\infty \text{ y } s(0)\to 0} \frac{1}{t} \log\left(\frac{s(t)}{s(0)}\right), \tag{8.218}$$

que nos da el promedio de la divergencia de las trayectorias y es, por tanto, una medida de la sensibilidad a condiciones iniciales de la solución $x_0(t)$.

Cuando los coeficientes que aparecen en la ecuación (8.216) son constantes, cuestión que sólo ocurre si la ecuación inicial es lineal o la solución $x_0(t)$ es estacionaria, entonces $\delta_i(t)$ será una combinación lineal de las exponenciales de los eigenvalores de la matriz definida por ((8.216)) multiplicados por el tiempo y como estos pueden ser complejos debemos entender que $\delta_i^2(t) = \delta_i(t)\delta_i^*(t)$. En este caso el límite ((8.218)) estará dominado por la mayor de las partes reales de los eigenvalores y será precisamente ese valor, vemos pues que si σ es positiva habrá al menos un eigenvalor con parte real positiva y, en torno a la solución $x_0(t)$, tendremos alta sensibilidad a condiciones iniciales. En el caso general la matriz definida por ((8.216)) será dependiente del tiempo y sumamente difícil de calcular, ya que requerimos la solución $x_0(t)$ para hacerlo, sin embargo σ seguirá siendo una medida de la separación exponencial de las soluciones cercanas a $x_0(t)$, por lo que si es positiva habrá sensibilidad a condiciones iniciales. A σ se le conoce como el coeficiente Liapunov y se han diseñado diversos procedimientos numéricos para calcularlo.

En realidad podemos tener un coeficiente de Liapunov por cada dimensión del espacio fase. Pensemos en una pequeña esfera centrada en un punto de la trayectoria $x_0(t)$, después de un tiempo pequeño, el flujo habrá deformado esta esfera en un elipsoide, esto debido a que no ocurre la misma expansión o contracción en todas direcciones. Infinitesimalmente los ejes de elipsoide corresponden a los eigenvectores de la matriz jacobiana definida en ((8.216)) y el coeficiente de expansión o contracción será el eigenvalor correspondiente, naturalmente esto cuando tenemos eigenvalores reales, en caso de tener eigenvalores complejos estos aparecerán en pares conjugados, el elipsoide tendrá una simetría circular por cada par de eigenvectores complejos y tendremos dos coeficientes iguales en la parte real de dichos eigenvalores. Los coeficientes o direcciones se ordenan, por convención, de mayor a menor y

definimos los coeficientes de Liapunov σ_i como el promedio de los coeficientes anteriores al transcurrir el tiempo, esto es

$$\sigma_k = \lim_{t\to\infty \text{ y } \delta_k(0)\to 0} \frac{1}{2t} \log\left(\frac{\delta_k^2(t)}{\delta_k^2(0)}\right), \tag{8.219}$$

donde δ_k es la separación en la dirección kaésima bajo la convención que establece que la primera dirección es la de mayor crecimiento, y así sucesivamente. Estos coeficientes nos dan una medida de la divergencia de las soluciones y, dado que lo más importante de estos coeficientes es tener un valor positivo, negativo o cero, se conoce como espectro de Liapunov a la configuración de los signos de dichos coeficientes. Si alguno de ellos es positivo tendremos alta sensibilidad a condiciones iniciales. El coeficiente σ definido en ((8.218)) resultará ser σ_1 ya que en el límite el coeficiente que representa la mayor divergencia será el que domine.

En el caso hamiltoniano el teorema de Liouville obliga a que la expansión en cierta dirección se compense con una contracción en otra, pues de lo contrario no se conservaría el volumen en el espacio fase, esto conduce a que la suma de los coeficientes de Liapunov para estos sistemas sea cero.

El cálculo de los coeficientes de Liapunov e incluso el del coeficiente σ o el espectro de Liapunov, resulta ser en general muy complicado, en especial si se intenta hacer aplicando la definición directamente. Se han desarrollado diversas técnicas para hacer estos cálculos, cada una con sus ventajas y limitaciones, algunas adecuadas al caso en que tenemos soluciones numéricas de ecuaciones diferenciales, otras al caso del análisis de datos experimentales, pero no es este el espacio para desarrollarlas y dejamos esto para los textos especializados.

Es importante señalar que la presencia de un coeficiente de Liapunov positivo indica alta sensibilidad a las condiciones iniciales, pero no necesariamente que se presente el fenómeno del caos, basta hacer ver que para las ecuaciones diferenciales lineales los coeficientes de Liapunov coinciden con la parte real de los eigenvalores de la matriz que define la ecuación, los que pueden ser positivos aunque sabemos que las ecuaciones lineales no puede tener soluciones caóticas. En todo caso la presencia de un coeficiente positivo es necesaria, pero no suficiente, para tener caos; sin embargo, cuando sabemos que la ecuación diferencial es no lineal, es muy probable que se tenga caos y bastará un pequeño análisis para eliminar otras posibilidades.

En el caso de los sistemas disipativos, contrario a lo que ocurre en los sistemas hamiltonianos, es frecuente encontrar sistemas con varios coeficientes negativos, uno o pocos positivos y ser la suma de los coeficientes negativa, cuando esto ocurre, en general tendremos que cierta región del espacio fase

se contrae, debido a los coeficientes negativos, a un objeto matemático de dimensión menor que la del espacio fase e inmerso en ese espacio; sobre este objeto invariante ante el flujo las trayectorias divergen, esto es, las trayectorias que comienzan cerca de dicho objeto tienden asintóticamente a él y las trayectorias sobre el objeto no salen de él y son caóticas. Estos objetos tienen una geometría sumamente compleja y se les conoce como atractores extraños. En el caso en que las trayectorias sobre el atractor invariante no fuesen caóticas, el objeto sería una variedad diferencial y lo llamaríamos un atractor regular. Naturalmente en el caso hamiltoniano no pueden existir ningún tipo de atractores debido a que el flujo preserva el volumen.

Hemos hecho un breve recorrido por el estudio del comportamiento de los sistemas mecánicos no integrables y del principal fenómeno que en ellos ocurre, el caos determinista. No hemos abarcado un sinnúmero de cuestiones importantes, ni dado indicación de lo mucho que hoy se investiga sobre estas cuestiones o de sus repercusiones en muchos campos de la ciencia, la tecnología e incluso las disciplinas filosóficas. No es la intención de un texto general de mecánica clásica el hacer esto, pero no podemos dejar de destacar que la elucidación de las propiedades de los sistemas no integrables ha llevado a la apertura de toda una nueva área del conocimiento científico, lo que, a su vez, ha hecho renacer a la venerable ciencia de la mecánica clásica como un campo de investigación dinámico y contemporáneo.

Índice alfabético

Constante de movimiento, 28

Diagrama fase, 29
    método gráfico, 29

Integral de movimiento, 28

Potencial
    en una dimensión, 27